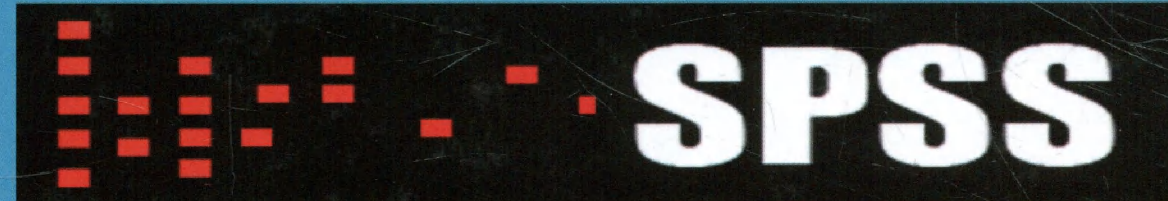
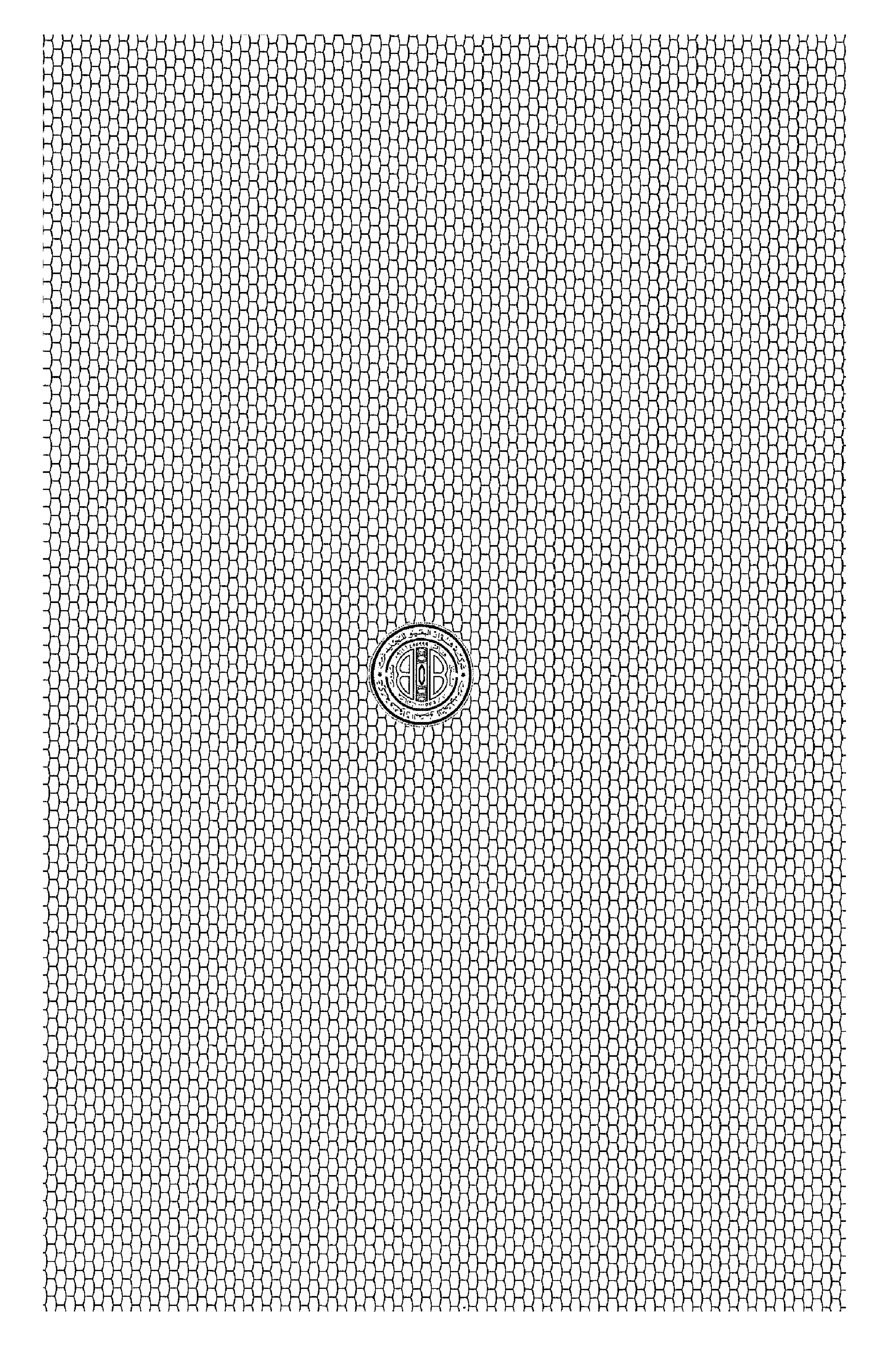
الإحصاء التحليلي

مع تطبيقات برمجية



الدكتور
نبيل جمعة صالح النجار







الإحصـــاء التحليلي

مع تطبيقات برمجية SPSS

الإحصاء التحليلي

مع تطبيقات برمجية SPSS

د. نبيل جمعه صالح النجار

دكتوراه الفلسفة في التربية

قسم القياس والتقويم



رقم التصنيف : 519.50.285
المؤلف ومن هو في حكمه : النجار، نبيل جمعة صالح.
عنوان الكتاب : الاحصاء التحليلي مع تطبيقات برمجية SPSS.
رقم الإيداع : 2014/6/2724
الواصفات : /الاحصاء//الحواسيب/
بيانات الناشر : عمان — دار ومكتبة الحامد للنشر والتوزيع
يتحمل المؤلف كامل المسؤولية القانونية عن محتوى مصنفه ولا يعبّر هذا المصنف عن رأي دائرة المكتبة الوطنية أو أي جهة حكومية أخرى.
(ردمك) ISBN 978-9957-32-873-3

تم إعداد بيانات الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية.

الطبعة الأولى 2015-1436هـ



دار الحامد للنشر والتوزيع

الأردن - عمان - شفا بدران - شارع العرب مقابل جامعة العلوم التطبيقية
هاتف: 5231081 6 962+ فاكس: 5235594 6 962+
ص.ب . (366) الرمز البريدي: (11941) عمان - الأردن
www.daralhamed.net
E-mail : daralhamed@yahoo.com

المحتويات

الفصل الخامس

الفصل السادس

الفصل السابع

الفصل الحادي عشر
التحليل العاملي
Factor Analysis

الفصل الثاني عشر
الاحصاءات اللامعلمية
NONPARAMETRIC STATISTICS

مُقَدِّمَة

الحمــد لله رب العالمين والصــلاة والســلام على ســيدنا محمد صلى الله عليه وسلم خير الأنبياء المرسلين.

لقــد أطلــق علــى هــذا العصر عصر المعلوماتية، إن هذا العصر هو عصر تكنولوجيا المعلومات، فقد أصبحت المعلومات تشكل الثورة الحقيقية التي سادت في الحياة العصرية ودخلت إلى بيوتنا وأعمالنا بشكل أو بآخر شئنا ذلك أم أبينا، لذا يجب علينا أن نساير ونواكب التقدم في شتى المجالات والميادين.

وبمــا أن الإحصاء علم مهم للكثير من الطلاب والباحثين فقد تم تغطية جميع فصول هذا الكتاب بتطبيق ما ورد فيها من معلومات نظرية من خلال التمارين العملية وعن طريق استخدام برمجية SPSS.

لقد راعى المؤلف أن يكون هذا الكتاب متوافقاً مع خطة منهاج **الإحصاء التحليلي للبحوث التربوية**" المقرر في كلية العلوم التربوية في جامعة مؤته، ومع المستوى المطلوب لمادة "**الإحصاء التحليلي**" و"**الاحصاء المتقدم**" على مستوى الجامعات الحكومية والخاصة عموماً، مع الأخذ بعين الاعتبار أن يكون مناسباً لجميع من لديه حب المعرفة بعلم الإحصاء التحليلي وتطبيقاته وفوائده وبطريقة سهلة وبسيطة.

ويــأتي هــذا الكــتاب كأحد الوسائل التي توفر لطلبة الدراسات العليا على مستوى الجامعات فرصة تمكنه من تعلم أساسيات مادة **الإحصاء التحليلي**". وذلك من خلال ما يحويه مــن مواضيع جديدة وهامة في مفاهيم ومبادئ **الإحصاء التحليلي** وتطبيقاته باستخدام برمجية **SPSS**.

د. نبيــل جمعــه النجـــار

Email: nabilnajjar@yahoo.com
Mobile: 0777757837
Mobile: 0785623442
Mobile: 0798011404

الفصل الأول

مفاهيم أساسية في الإحصاء

Statistical Concepts

1-1 المفاهيم الإحصائية Statistical Concepts

مفهوم الاحصاء، الرمـوز الإحصائية، الـمـتـغـيرات.

2-1 مقاييس النزعة المركزية Measures of Central Tendency

الوسط الحسابي، الوسـيــط ، الـمـنــوال ، الـمئينـات.

3-1 مقاييس التشتت Measure Dispersion or Variation

المــدى، الانحـــراف المتوســـط، التـــباين، الانحـــراف المعياري، معامل الاختـلاف.

4-1 الارتبـــاط Correlation والانـحــــــدار Regression

أنواع الارتباط ، قياس الارتباط.

الانحدار ومفهـومــة، الانحدار الخطي البسيط، معادلـة خــط الانحدار.

5-1 استخدام برمجية SPSS في حـل المــسائل.

6-1 تـمـاريـــن Exercise.

الفصل الأول

مفاهيم أساسية في الإحصاء
Statistical Concepts

1-1 المفاهيم الإحصائية Statistical Concepts

مفهوم الإحصاء Definition of Statistics

مجموعة الطرق العلمية التي تعنى بجمع وتصنيف وتبويب وتفسير وتلخيص وتقييم البيانات والخروج منها باستنتاجات حول المجتمع من خلال اعتماد جزء صغير من المجتمع (العينة).

الإحصاء الوصفي Descriptive Statistics

جمع المعطيات وتحليلها ووصفها وإظهارها بصيغة مفهومة وذات مدلول والتعامل مع المعطيات الإحصائية من دون تعميم، وعرضها عن طريق الجداول والرسوم البيانية وغيرها.

الإحصاء الاستدلالي Inferential Statistics

تحليل وتفسير وتقدير واستخلاص الاستنتاجات بالاعتماد على عينة من المجتمع للتوصل إلى قرارات تخص المجتمع ويتعامل مع التعميم والتنبؤ.

المعطيات الإحصائية Statistical Data

البيانات والمعلومات الإحصائية المتعلقة بالظواهر الإدارية والاجتماعية والتربوية وتختلف المعطيات من حيث نوعها وطبيعتها باختلاف الظاهرة المطلوب قياسها وباختلاف منهجية البحث والأدوات الإحصائية المستخدمة.

البيانات Data: مشاهدات، علامات، مقادير، يعبر عنها بأرقام.

البيانات غير المبوبة Ungrouped Data: البيانات الأولية أو الأصلية التي جمعت و لم تبوب.

البـــيانات المـــبوبة **Grouped Data**: البيانات التي بوبت وفرغت في جدول توزيع تكراري.

المعطيات الكمية Quantitative Data

تصف الظاهرة بشكل رقمي عن ظاهرة معينة، مثل علامة الطالب وسعر السلعة.

المعطيات النوعية Qualitative Data

تصف الظاهرة المعنية بشكل غير رقمي عن ظاهرة معينة، مثل الجنس، اللون.

التوزيع Distribution: مجموعة مشاهدات مهما كان عددها.

المجتمع Population: مجتمع بيانات أو مشاهدات أو علامات يحدد هويته الباحث.

مجتمع العينة Sample Population: المجتمع الذي تؤخذ منه العينة.

مجـــتمع الهدف Target Population: المجتمع الذي ستعمم عليه نتائج الدراسة التي أجريت على مجتمع العينة.

العينة Sample: مجموعة جزئية من المجتمع.

المؤشـــر Index: تدل على جميع مقاييس النزعة المركزية والتشتت والعلاقة (الارتباط) سواء محسوبة لعينات أو لمجتمعات.

مؤشر عـــينة (إحصائي) Statistic:ويستخدم للعينات، مثل الوسط الحسابي لعينة X'.

مؤشـــر مجتمع (مـــعـــلم) Parameter: ويستخدم للمجتمع، مثل الوسط الحسابي لمجتمع μ.

لكل **توزيع خصائص** (توصيف) هي: مقاييس النزعة المركزية.

*** الرمـــوز الإحصائية Statistical Symbols**

في الإحصاء الاستدلالي تستخدم إحصاءات العينة كتقديرات لمعالم المجتمع المناظر.

رموز العينة	المعنى	رموز المجتمع	المعنى
X'	متوسط عينة	μ_x	متوسط مجتمع.
S^2_x	تباين عينة	σ^2_x	تباين مجتمع.

الرمز	المعنى	الرمز	المعنى
S_x	انحراف معياري عينة	σ_x	انحراف معياري لمجتمع.
r_{xy}	معامل ارتباط لعينة	ρ_{xy}	معامل ارتباط لمجتمع.
P	نسبة عينة	π	نسبة المجتمع.
n	عدد مشاهدات العينة	N	عدد مشاهدات المجتمع.

الرمز	المعنى	English
n	عدد المشاهدات.	Number of Scores
Σx	مجموع المشاهدات.	Sum of x
M_d	الوسيط.	Median
$\overline{X}$	الوسط الحسابي للعينة.	Sample Mean of X's
μ	الوسط الحسابي للمجتمع.	Mu; Population Mean
X, Y	علامة خام.	Raw Score
f	التكرار.	Frequency
cf	التكرار التراكمي.	Cumulative Frequency
$rel.f$	التكرار النسبي.	Relative Frequency
df	درجات الحرية.	Degree Freedom
S_x	انحراف معياري عينة.	Sample Standard Deviation
S^2_x	تباين عينة.	Sample Variance
σ_x	انحراف معياري لمجتمع.	Population Standard Deviation
σ^2_x	تباين مجتمع.	Population Variance
ρ	معامل ارتباط مجتمع.	Population Correlation Coefficient
r	معامل ارتباط عينة.	Sample Correlation Coefficient
Z	العلامة المعيارية.	Standard Score
$\sigma_{x'}$	الخطأ المعياري للوسط.	Standard Error of Mean
D	الفرق بين البيانات. (X-Y)	Difference Score
r	معامل ارتباط بيرسون.	Pearson Correlation Coefficient
r_s	معامل ارتباط سبيرمان.	Sperman Correlation Coefficient

r_{pb} ر بوينت بايسيريال. Point-biserial Correlation Coefficient

$Y=ax+b$ معادلة خط الانحدار Regression Line Equation

b ميل خط الانحدار. Slope of the Regression Line

a نقطة تقاطع خط الانحدار. r^2 Y-intercept of the Regression Line

معامل التحديد. $\overline{D}$ Coefficient of Determination

الوسط الحسابي للفروق. Mean of Difference Scores ($\Sigma D/n$)

*** المتغيرات Variables**

المتغيرات Variables ما إحصائية أو عشوائية فالمتغير الإحصائي يمثل القيم التي تأخذها ظاهرة ما، والمتغير العشوائي عبارة عن ظاهرة نوعية أو كمية لا يمكن التنبؤ بها بشكل مسبق.

المتغير Variable: ظاهرة تظهر اختلافات بين قيمها، إذا اختلفت الخاصية عند أفراد مجموعة معينة كماً أو نوعاً نقول بأنها هي المتغير، إذا كان الأفراد متساوين كماً أو متشابهين نوعاً بالنسبة لخاصية معينة تكون هي الثابت.

*** تصنف المتغيرات حسب طبيعة المعلومات التي يؤديها القياس إلى:**

1- المتغيرات الاسمية Nominal Variables

المتغيرات النوعية التي لها عدد فئات محدد من دون أي وزن لهذه الفئات ولا يوجد أفضلية لأحدها على الآخر، وتستخدم لغايات التصنيف فقط.

مثال: متغير الجنس ويصنف فيه المجتمع إلى فئتين هما الذكور والإناث فلو رمزنا للذكور بالرقم (1) والإناث بالرقم (2) فالأرقام ليس لها معنى حقيقي ولا يمكن إجراء العمليات الحسابية عليها.

مثال: إذا قسم الأفراد حسب الطول إلى طويل وقصير.

مثال: أوجه قطعة النقد وهي صورة وكتابة.

2- المتغيرات الترتيبية Ordinal Variables

متغير نوعي ذو عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً، ولا يمكن تحديد الفروق بدقة بين القيم المختلفة. مثال: كبير وسط صغير.

مثال: A أكبر من B ولكن لا نستطيع معرفة كم يكبر A عن B.

مثال: إذا كانت علامة جمال في مادة اللغة العربية أكثر من علامة فادي، وأن علامة فادي أكثر من علامة نبيل، فإننا نعرف هنا ترتيب الأفراد فقط.

3- المتغيرات الفئوية Interval Variables

المتغيرات الكمية التي يمكن إجراء العمليات الحسابية عليها وذلك دون أن تتأثر المسافة النسبية بين قيمها، ويميز هذا المتغير من خلال قيمة الصفر التي لا تعني عدم توافر تلك الصفة.

مثال: إذا كانت علامة جمال في مادة اللغة العربية (60) وعلامة أيمن (50)، وعلامة محمد (40)، فإننا نعرف الترتيب، ونعرف كم تزيد علامة كل طالب عن الآخر.

مثال: إذا حصل محمد على علامة (صفر) في اختبار اللغة الإنجليزية فهذا لا يعني أن محمد لا يعرف شيئاً في اللغة الإنجليزية.

مثال: إذا كانت درجة الحرارة تساوي (صفر) فهذا لا يعني عدم وجود درجة حرارة.

4- المتغيرات النسبية Ratio Variables

متغيرات كمية ليس لها فئات محددة وهي تشبه المتغيرات الفئوية ولكن الصفر هنا يمثل عدم توفر الصفة، ومثال له المتغيرات الزمنية.

مثال: علامات الطلاب في مساق معين، أوزان الطلاب.

المقياس	الاستخدام
1.- الاسمي Nominal	تصنيف
2.- الترتيبي Ordinal	تصنيف + ترتيب
3.- الفئوي Interval	تصنيف + ترتيب + مسافة + صفر افتراضي
4.- النسبي Ratio	تصنيف + ترتيب + مسافة + صفر مطلق

*** تصنف المتغيرات حسب وجود علاقة بين متغيرين إلى:**

1. **المــتغير المــستقل Independent Variable**: وهو المتغير الذي يخضع لسيطرة الإحصائي أو الباحث.

2. **المــتغير التابع Dependent Variable**: وهو المتغير الذي تتنبأ بقيمته من خلال معرفتنا لقيم المتغير المستقل.

مثال: إذا أراد مدرس أن يبحث عن أثر **عدد ساعات الدراسة** على **تحصيل الطالب** في مبحث معين، حدد المتغير المستقل والمتغير التابع؟

المتغير المستقل: عدد ساعات الدراسة. المتغير التابع: تحصيل الطالب.

مثال: أثر **الغياب** على **تحصيل الطالب**.

المتغير المستقل: الغياب. المتغير التابع: تحصيل الطالب.

2-1 مقاييس النزعة المركزية
Measures of Central Tendency

مقدمة:

لأي بــيانات إحــصائية هناك خواص تساعد على إعطاء فكرة ومدلول عن وضع هذه البيانات ومن هذه الخصائص ما يلي:

– **إحصائيات النزعة المركزية** وتسمى بالمتوسطات واهم هذه المتوسطات هي: الوسط الحسابي والوسيط والمنوال، ونتمكن بواسطتها من تحديد موقع النقطة التي تتمحور حولها كافة القيم، ولكل من هذه الإحصائيات مزايا وعيوب، ويمكن استخدام أي منها بناء على عدة امور منها:

- شكل التوزيع: هل هو معتدل أم ملتو.
- مستوى القياس: هل هو اسمي، رتبي، فئوي، نسبي.

– **إحــصائيات التــشتت**: ويقــصد بها حالة الانتشار التي تكون عليها البيانات حول المتوسط، وأهم هذه المقاييس المدى والتباين والانحراف المعياري.

1- الوسط الحسابي ($\overline{X}$) Arithmetic Mean

الوسط الحسابي هو معدل المشاهدات في التوزيع.

الوسط الحسابي هو مجموع قيم المشاهدات على عددها.

* الوسط الحسابي للبيانات غير المبوبة Ungrouped Data

الوسط الحسابي $\bar{س} = \frac{\text{مجموع المشاهدات}}{\text{عددها}}$ $\overline{X} = \frac{\Sigma X_i}{n}$

حيث أن:

X : رمز الوسط الحسابي للعينة ويقرأ اكس بار.

Σ : رمز المجموع ويقرأ سيجما.

X_i : رمز المشاهدة.

n : عدد المشاهدات في العينة.

* الوسط الحسابي للبيانات المبوبة Grouped Data

$$\bar{س} = \frac{\sum س_ي * ت_ي}{\sum ت_ي}$$

$$\overline{X} = \frac{\Sigma X_i * F_i}{\Sigma F_i}$$

F_i : التكرار. X_i : مراكز الفئات.

* الدلالات الإحصائية للوسط الحسابي Statistical Significance for Mean

- كلما ارتفعت قيمة الوسط الحسابي للعلامات دل ذلك على أداء أفضل، بشرط أن لا تكون هناك قيم متطرفة عالية أدت إلى ارتفاع الوسط الحسابي .
- كلما كانت العلامات موزعة على جانبي وسطها الحسابي بشكل متماثل ومتساو كان التوزيع معتدلاً وكاشفاً عن الفروق بين الطلاب بصورة أفضل.

* خصائص الوسط الحسابي Mean's Characteristics

- يعتمد على جميع المشاهدات.
- سهل الفهم والتفسير، ويتم حسابه بسهولة وسرعة.
- يتأثر بالتحويلات الخطية، ولا يتأثر باختلاف العينات في المجتمع.
- مقياس النزعة المفضل عند الحديث عن الإحصاء الاستدلالي.
- مجموع انحرافات القيم عن وسطها الحسابي يساوي صفراً.

* عيوب الوسط الحسابي Mean's Deficiencies

- لا يمكن قياسه بالطرق البيانية.
- قابليته للتأثر بعدد قليل من المشاهدات المتطرفة.
- لا يمكن حسابه في التوزيع ذي الفئات المفتوحة.

2- الوسيط Median (M_d)

المشاهدة التي تقسم التوزيع إلى نصفين بحيث يكون فوقها 50% من المشاهدات ودونها 50% من المشاهدات .

الوسيط: المشاهدة التي تقع في منتصف التوزيع. وهو المئين 50 (P50).

لحساب الوسيط:

- نعمل على ترتيب البيانات تصاعدياً أو تنازلياً.
- إذا كان عدد البيانات فردياً تكون المشاهدة التي ترتيبها 2/(1+n) هي الوسيط.
- إذا كان عدد البيانات زوجياً يكون معدل المشاهدتين اللتان ترتيبهما (n/2) و 1+(n/2) هو الوسيط.

* الوسيط للبيانات المبوبة Grouped data

$$\text{و} = \text{الحد الأدنى الفعلي لفئة الوسيط} + \frac{(\text{ت}_{\text{و}} - \text{ت}_{\text{ص}}\ \text{للفئة السابقة})}{\text{ت}_{\text{ر}}\ \text{لفئة الوسيط}} * \text{ط}$$

$\text{ت}_{\text{و}}$: التكرار التراكمي للوسيط.

تص : التكرار التراكمي الصاعد للفئة السابقة لفئة الوسيط.

تر : التكرار العادي لفئة الوسيط.

ط : طول الفئة.

$$M_d = L + \frac{(\Sigma fi/2) - f_1}{f_2 - f_1} * H$$

حيث أن:

L : الحد الأدنى الفعلي لفئة الوسيط. H :طول الفئة. $(\Sigma F_i /2)$: قيمة موقع الوسيط.

f1 : التكرار المتجمع السابق لموقع الوسيط f2 :التكرار المتجمع اللاحق لموقع الوسيط

خصائص الوسيط Median's Characteristics

- يمكن احتسابه للجداول المفتوحة.
- يمكن احتسابه في حالة فقدان بعض القيم شرط ان يكون ترتيبها معروفاً.
- سهولة احتسابه، يمكن إيجاده بيانياً.
- لا يتأثر بالقيم المتطرفة، موقعه يتوسط البيانات.
- لا يعتمد على قيم البيانات وإنما يعتمد على موقعها.
- يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة).

عيوب الوسيط Median's Deficiencies

- حساس للقيم الوسيطة.
- إذا كان عدد المشاهدات قليل فالوسيط قد لا يعبر بصورة واضحة صحيحة عن مركز تجمع المشاهدات.

3- المنوال (M_o) Mode:

العلامة أو مركز الفئة أو الصفة التي تقابل أعلى تكرار في التوزيع.

القيمة الأكثر تكراراً أو الظاهرة الأكثر شيوعاً.

* المنوال للبيانات غير المبوبة **Ungrouped Data**

- ليس من الضروري وجود منوال للبيانات.

* المنوال للبيانات المبوبة **Grouped Data**

1- طريقة الفروق (بيرسون):

- تحديد الفئة المنوالية: وهي الفئة التي يقابلها أكبر تكرار.
- استخدام الصيغة التالية لحساب قيمة المنوال

$$M_o = L + \frac{d_1}{d_1 + d_2} * H$$

حيث ان: L : الحد الأدنى الفعلي للفئة المنوالية. H : طول الفئة.

d_1 : تكرار الفئة المنوالية – تكرار الفئة السابقة

d_2 : تكرار الفئة المنوالية – تكرار الفئة اللاحقة

2- طريقة الرافعة

$$M_o = L + \frac{d_a}{d_a + d_b} * H$$

حيث ان: L : الحد الأدنى الفعلي للفئة المنوالية. d_a : تكرار الفئة اللاحقة.

H : طول الفئة. d_b : تكرار الفئة السابقة.

المنوال = الحد الأدنى الفعلي للفئة المتوالية + (تكرار الفئة اللاحقة ÷ (تكرار الفئة اللاحقة+ تكرار الفئة السابقة)) × طول الفئة

خصائص المنوال Mode's Characteristics

- محدود الاستعمال ، و يتأثر كثيراً بحجم العينة.
- الإحـــصائي الوحـــيد الذي يمكن استعماله عندما تكون البيانات الإحصائية بمستوى القياس الاسمي.

- يمكن استعماله عندما تكون البيانات رتبيه أو فئوية أو نسبة.
- لا يعتمد على جميع قيم البيانات وإنما يعتمد على القيم المتكررة.
- يتأثر بطول الفئة في التوزيع، ويمكن إيجاده بيانياً.
- يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة).
- لا يتأثر بالقيم المتطرفة.
- قد يكون للتوزيع أكثر من منوال.
- لا يدخل كثيراً في تحليلات إحصائية خارج نطاق وصف البيانات.
- يمكن احتسابه للجداول المفتوحة.
- اقل تعبيراً كمتوسط عندما تكون القيم منتشرة على مدى واسع.
- يمكن أن يكون للبيانات أكثر من منوال واحد.

مثال 4-11: البيانات التالية: 5، 10، 27، 32، 18، 10، 40، 32 ليس من الضروري وجود منوال للبيانات.

- المنوال يمكن أن يحسب للمتغيرات الكمية والمتغيرات النوعية.

4- المئينات (P_i) Percentiles :

مقياس يتم بموجبه تقسيم البيانات إلى 100 جزء متساوي وبالتالي يوجد 99 مئين. وهو قيمة معينة ضمن التوزيع تسبقها أو تليها نسبة مئوية معينة من المشاهدات الداخلة فيه.

المئين ك: المشاهدة التي يقل عنها أو يساويها ك% من المشاهدات، ويرمز للمئين ك بالرمز مك، إن قيمة المئين i تحسب عن طريق الصيغة التالية:

$$P_i = L + \frac{[(i \sum f_i) / 100] - f_1}{f_2 - f_1} * H$$

*** إيجاد المئين للبيانات غير المبوبة Ungrouped Data**

تعريف: **الرتبة المئينية** لمشاهدة ما هي النسبة المئوية للتكرار التراكمي المقابل لتلك المشاهدة بالنسبة إلى مجموع التكرارات.

إيجاد رتبة المئين المطلوب = (ترتيب المئين/100) * (عدد البيانات+1)

رتبة المئين= ك * (n+1) / 100	Rank $P_k = k(n+1)/100$

* إيجاد المئين للبيانات المبوبةGrouped Data

إيجاد موقع أي مئين يتم تحديده كالآتي: $(i \sum f_i) / 100$

موقع المئين ثمانين P80 هو القيمة $(80 \sum f_i) / 100$

تحديد التكرار الصاعد المقابل للمئين المطلوب.

تحديد فئة المئين المطلوب: وهي أول فئة تقابل تكرار صاعد يساوي أو يزيد عن التكرار الصاعد للمئين.

المئين ك= الحد الأدنى الفعلي لفئته+ (ك/100)*مجموع(تي– تكرار صاعد للفئة السابقة)* ط
تكرار عادي لفئة المئين

إن قيمة المئين i تحسب عن طريق الصيغة التالية:

$$P_i = L + \frac{[(i \sum f_i) / 100] - f_1}{f_2 - f_1} * H$$

3-1 مقاييس التشتت Measure Dispersion or Variation

تستخدم مقاييس التشتت لقياس انتشار قيم المشاهدات حول نقطة التركز وهي الوسط الحسابي، وإن المقصود بالتشتت أو الاختلاف هو التباعد الموجود بين قيم المشاهدات التابعة لمتغير ما عن وسطها الحسابي.

إذا كان لدينا مجموعتان من البيانات لها نفس الوسط الحسابي فمن الممكن أن تكونا مختلفتين في انتشارهما حول الوسط الحسابي.

كلما كبرت قيم مقاييس التشتت دل ذلك على درجة كبيرة من الاختلاف بين قيم البيانات، وكلما صغرت قيم مقاييس التشتت دل ذلك على درجة قليلة من الاختلاف بين قيم البيانات.

لذلك هذه المقاييس تعطي فكرة عن مدى تجانس أو اختلاف البيانات عن مركزها ويدل ذلك على درجة انتشارها.

وهناك عدة مقاييس للتشتت وهي:

1. المـــــــدى Range (R)
2. الانحراف المتوسط Mean Deviation (MD)
3. التـــبـــايـــن Variance (σ^2)
4. الانحراف المعياري Standard Deviation (σ)
5. معامل الاختلاف Coefficient Variation (CV)

1- المدى Range (R)

يعـــرف المـــدى لمجمـــوعة من البيانات على انه الفرق بين اكبر قيمة واصغر قيمة لتلك المجموعة، ويرمز له بالرمز R.

*** المدى للبيانات غير المبوبة Ungrouped Data**

المدى = اكبر قيمة – اصغر قيمة Range = Highest score - Lowest score

*** المدى للبيانات المبوبة Grouped Data**

بمـــا أن اصغر قيمة واكبر قيمة مجهولة في حالة المعطيات المبوبة فان قيمة المدى التقديرية هي:

المدى للتوزيع = مـــركـــز الفئة الأخيرة – مركز الفئة الأولى
= الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى
= الحد الأعلى الفعلي للفئة الأخيرة – الحد الأدنى الفعلي للفئة الأولى

*** عيوب المدى Range's Deficiencies**

1. يتأثر بالقيم الشاذة والمتطرفة.
2. لا يستخدم في حالة المعطيات المبوبة التي تتضمن فئات مفتوحة.

* مزايا المدى **Range's Characteristics**

1. سهولة الفهم.
2. سهولة حسابه.
3. كثرة استخدامه في الأوساط العامة.

2- الانحراف المتوسط **Mean Deviation (MD)**

الوسط الحسابي للقيمة المطلقة لانحراف تلك القيم عن وسطها الحسابي.

معدل انحراف المشاهدات في التوزيع عن وسطه الحسابي.

* الانحراف المتوسط للبيانات غير المبوبة **Ungrouped Data**

$$\text{أ.م} = \frac{\sum |\text{س}_\text{ر} - \bar{\text{س}}|}{\text{ن}}$$

$$MD = \frac{\Sigma |x_i - \bar{X}|}{n}$$

حيث: M_D ، أ.م : الانحراف المتوسط.

Xi ، $\text{س}_\text{ر}$: المشاهدات. $\bar{X}$ ، $\bar{\text{س}}$: الوسط الحسابي.

* الانحراف المتوسط للبيانات المبوبة **Grouped Data**

بافتـــراض أن كافة القيم الواقعة ضمن الفئات هي عند مركز هذه الفئات فإن الانحراف المتوســـط عـــبارة عـــن مجموع انحرافات قيم مراكز هذه الفئات عن الوسط الحسابي مضروبة بتكراراتها ثم قسمة الناتج على مجموع التكرارات.

$$\text{أ.م} = \frac{\sum |\text{س}_\text{ر} - \bar{\text{س}}| * \text{ت}_\text{ر}}{\sum \text{ت}_\text{ر}}$$

$$MD = \frac{\Sigma D_i * f_i}{\Sigma f_i}$$

حيث:

| س ر − س̄ | ، D_i : قيم الانحرافات المطلقة عن الوسط الحسابي $|x_i - \overline{x}|$

X_i : مراكز الفئات. ت ر ، F_i : التكرار.

*** عيوب الانحراف المتوسط Mean Deviation Deficiencies**

1. نادر الاستخدام بسبب كون عملية احتسابه تعتمد على القيم المطلقة والتي تهمل الإشارة.
2. عدم إمكانية استخدامه مع الجداول التكرارية ذات الفئات المفتوحة.

*** مزايا الانحراف المتوسط Mean Deviation Characteristics**

1. في احتسابه يتم شمول كل القيم المطلوب تقدير قيمة تشتتها.

3- التباين Variance (σ^2)

هو الوسط الحسابي لمربعات انحرافات القيم Xi عن وسطها الحسابي $\overline{X}$.

القيمة المرتفعة للتباين تعني أن الأشياء متباينة، متباعدة، متناثرة، غير متجانسة.

القيمة المنخفضة للتباين تعني أن الأشياء غير متباينة، متقاربة، متجانسة.

التباين للبيانات غير المبوبة Ungrouped Data

$$ع^2 = \frac{\sum (س_ر - \bar{س})^2}{ن}$$

$$ع^2 = \frac{\sum (س_ر^2) - \bar{س}^2}{ن}$$

$$\sigma^2 = \frac{\Sigma(x_i - \overline{X})^2}{n}$$

$$\sigma^2 = \frac{\Sigma x_i^2 - n\overline{X}^2}{n}$$

$$\sigma^2 = \frac{\Sigma x_i^2 - [(\Sigma x)^2 / n]}{n}$$

* التباين للبيانات المبوبة **Grouped Data**

$$\sigma^2 = \frac{\Sigma(x_i - \overline{X})^2 * f_i}{\Sigma fi}$$

$$\sigma^2 = \frac{\Sigma X_i^2 * fi - n * \overline{X}^2}{\Sigma fi}$$

4- الانحراف المعياري (σ) Standard Deviation

الانحراف المعياري من أكثر مقاييس التشتت شيوعاً وأهمية، وهو الجذر التربيعي لمتوسط مـــربعات انحـــرافات القـــيم عن متوسطها الحسابي، وإن الانحراف المعياري هو الجذر التربيعي للتباين.

لحساب الانحراف المعياري:

- إيجاد المتوسط الحسابي للمشاهدات.
- إيجاد انحرافات القيم المختلفة عن المتوسط الحسابي ومن ثم تربيعها، ثم جمعها.
- إيجاد متوسط مربعات انحرافات القيم عن متوسطها الحسابي.
- إيجاد الجذر التربيعي لمتوسط مربعات انحرافات القيم عن متوسطها الحسابي.

القانون الأساسي للانحراف المعياري:

$$\sigma = \sqrt{\Sigma(X - \overline{X})^2 / n} = \sqrt{\sigma^2}$$

الانحراف المعياري = الجذر التربيعي للتباين

$$\sigma = \sqrt{\frac{\text{مج}(\text{س} - \overline{\text{س}})^2}{\text{ن}}} = \sqrt{\sigma^2}$$

الشكل العملي لقانون الانحراف المعياري

تعـــتمد هذه الصيغة على القيم الأصلية، ولا يدخل المتوسط أو الانحرافات عن الوسط الحسابي في حساها بشكل مباشر، وتكون المعادلة على الشكل التالي:

$$\sigma = \sqrt{\left(\frac{\Sigma x^2}{n}\right) - \left(\frac{\Sigma x}{n}\right)^2}$$

$$\sigma = \sqrt{\left(\Sigma x_i^2 - [(\Sigma x)^2 / n]\right) / n}$$

حيث أن: X: المشاهدات. n: عدد المشاهدات.

* خصائص الانحراف المعياري Standard Deviation Characteristics

1- يعتبر من أكثر مقاييس التشتت استخداماً وأهمية.

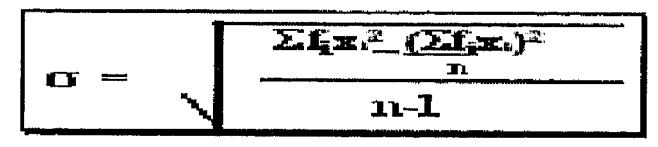
2- يتأثر بالتحويلات الخطية.

3- يعتمد في حسابه على جميع المشاهدات.

4- في التوزيعات القريبة من التوزيع الطبيعي نلاحظ أن:

- 68.27% من البيانات تقع في مدى انحراف معياري واحد عن الوسط.
- 95.45% من البيانات تقع في مدى انحرافين معياريين عن الوسط.
- 99.73% من البيانات تقع في مدى ثلاثة انحرافات معيارية عن الوسط.

* الانحراف المعياري للبيانات المبوبة Grouped Data



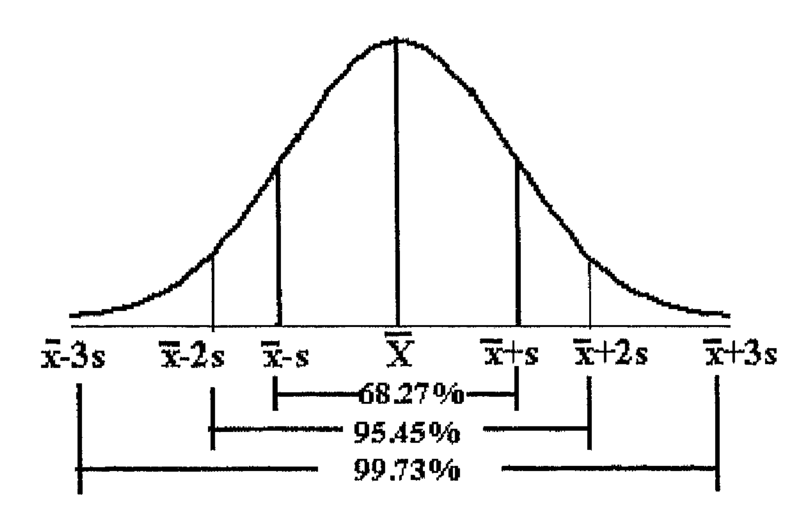
$$\sigma = \sqrt{\frac{\Sigma f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{n}}{n-1}}$$

حيث:

Xi: مركز الفئة. ، fi: التكرار. ، n: عدد المشاهدات.

* تفسير الانحراف المعياري Standard Deviation Explanation

يستخدم الانحراف المعياري بشكل عام كأحد مقاييس التشتت، وذلك لمعرفة عدد المشاهدات ونسبتها والتي تقع على بعد معين من الوسط الحسابي $\overline{X}$ حيث انه وباستخدام الصيغة التجريبية (في التوزيعات المتماثلة والطبيعية) نلاحظ أن نسب المشاهدات تكون كما هو موضح بالشكل(1).



شكل(1): الصيغة التجريبية للتوزيع المتماثل.

في الـــتوزيعات الطبيعـــية نتوقع أن 68.27% من القيم تقع على بعد انحراف معياري حـــول الوسط، وأن 95.45% من القيم تقع على بعد انحرافين معياريين حول الوسط، وأن 99.73% من القيم تقع على بعد ثلاث انحرافات معيارية حول الوسط، أما في التوزيعات غير المتماثلة نتوقع الحصول على نسب مختلفة عن تلك.

ولتوضيح عملية استخدام الانحراف المعياري لتفسير عينة ما، يجب إيجاد الفترات $\bar{X}\pm s$ ، $\bar{X}\pm 2s$ ، $\bar{X}\pm 3s$ ثم تحديـــد عدد البيانات ثم نسبتها، وبعد ذلك مقارنتها مع النسب المعيارية [68.27%, 95.45%, 99.73%] لـــتحديد مـــا إذا كان توزيع المشاهدات طبيعي والتشتت معـــتدل، أمـــا لمقارنة عينتين فيمكن مقارنة النسب ببعضها والحكم بعد ذلك أن كان تشتت إحدى العينتين أكبر أو أقل من العينة الأخرى.

5- معامل الاختلاف Coefficient of Variation (CV)

إن معامـــل الاختلاف يفضل عن الانحراف المعياري وغيره من مقاييس التشتت لمقارنة تشتت البيانات بين عدة مجموعات من البيانات.

إن معامل الاختلاف يعطي نسبة الانحراف المعياري إلى الوسط الحسابي، وبما أن معامل الاخـــتلاف هـــو مقياس لقياس التغير النسبي على شكل نسبة مئوية، لذلك معامل التغيير يمكن اســـتخدامه لمقارنة التشتت داخل عدة مجموعات من البيانات حتى وان كانت وحدات القياس لهذه المجموعات مختلفة.

ويحسب من خلال المقارنة بين الوسط الحسابي والانحراف المعياري للبيانات حسب :

$$CV = \frac{Sx}{\bar{X}} * 100\% \quad \text{or} \quad CV = \frac{\sigma}{\mu} * 100\%$$

معامل الاختلاف = (الانحراف المعياري/الوسط الحسابي)* 100%

كلما كان معامل التغير كبير يعني ذلك وجود تباين واختلاف بين المشاهدات.

4-1 الارتباط Correlation والانحدار Regression

إن نظرية الارتباط تظهر قوة العلاقة بين متغيرين مع إمكانية تحديد نوع وقوة العلاقة بين الظواهر، كالعلاقة بين مستوى التعليم والأداء، والعلاقة بين معدل الثانوية العامة ومعدل الجامعة، والعلاقة بين المستوى الاقتصادي والتحصيل.

إن الهدف من تحليل الارتباط Correlation هو معرفة وجود علاقة بين متغيرين أو مجموعة من المتغيرات المستقلة Independent Variables (X_i) مع المتغير التابع Dependent Variable (Y_i) من عدم وجودها، وهناك عدة مقاييس لتحديد درجة العلاقة والارتباط بين المتغيرات.

الارتباط Correlation: علاقة بين متغيرين لمعرفة ما إذا كان تغير أحدهما مرتبطاً بتغير الآخر.

- **أنواع الارتباط Types of Correlation**

* **من حيث القـــوة Based on Strength**

1. **ارتباط تام** Complete Correlation: يتحدد متغير كلياً عن طريق متغير آخر.

   مثال) العلاقة بين مساحة الدائرة ونصف قطرها.

2. **ارتباط غير تام** Partial Correlation: يتأثر متغير معين بمتغير آخر.

   مثال) حجم الإنفاق ودخل الأسرة.

* **من حيث عدد المتغيرات Based on Number of Variables**

1. **ارتباط بسيط r**: Simple Correlation ارتباط بين متغيرين كميين فقط.

   مثال) العلاقة بين عمر الأب وعمر الطالب.

2. ارتــباط متعدد R: Multiple Correlation يدرس العلاقة بين أكثر من متغيرين، تكون العلاقة بين المتغير التابع Y وعدة متغيرات مستقلة X_i.

مثال) العلاقة بين حجم المبيعات وعدد الأسواق وعدد السكان.

*** قياس الارتباط Measures of Correlation**

أ- معامل ارتباط بيرسون الخطي Person Linear Correlation Coefficient

معامل ارتباط بيرسون يقيس قوة واتجاه العلاقة الخطية فقط بين متغيرين كميين.

ب- معامل ارتباط سبيرمان للرتب Spearman Correlation Coefficient

يــستخدم معامــل ارتباط سبيرمان ومعامل كندال تاو لقياس قوة الارتباط بين متغيرين **ترتيبيين** Ordinal.

الانحدار ومفهومة Regression Concept

في دراســة العلاقــة بين ظاهرتين أو أكثر إذا كان الهدف تحديد نوع وقوة العلاقة فإننا نــدرس الارتــباط Correlation، أما إذا كان الهدف دراسة العلاقة من خلال التمثيل البياني بأفــضل علاقــة اقتران ممكنة بالشكل $Yi = f(x)$ فإننا ندرس الانحدار Regression ويسمى المــستقيم أو المــنحنى الذي يمثل هذه الاقتران بمستقيم أو منحنى الانحدار، وهو من الأساليب الإحــصائية المستخدمة لتحديد التأثيرات بين المتغيرات المستقلة والمتغير التابع عن طريق معادلة الانحــدار للتنــبؤ بقيمة المتغير التابع بدلالة المتغيرات المستقلة، فإذا كان عدد المتغيرات المستقلة واحد فيسمى انحدار خطي بسيط Simple Linear Regression، أما إذا كان عدد المتغيرات المستقلة أكثر من واحد فيسمى انحدار متعدد Multiple Regression

الانحدار Regression:

للانحدار عدة تعريفات منها:

- الميل أو الانحدار نحو الوسط.
- إيجاد النموذج الذي يمثل العلاقة السببية بين متغيرين أو أكثر.

- يستخدم للتنبؤ بقيمة متغير عن طريق معرفة متغير آخر مرتبط به، مثل التنبؤ بالأرباح إذا عرفت المبيعات.
- العلاقـــة بين المتغيرات من خلال بناء معادلة تستخدم للتقدير والتنبؤ بقيمة المتغير التابع Y بدلالة متغير أو متغيرات مستقلة Xi ، مثل العلاقة بين الدخل والطلب.
- الانحـــدار الخطي Linear Regression: تكون معادلة التقدير عند عرضها بيانياً على شكل خط مستقيم.

التنبؤ:

تقدير بيانات غير معروفة مبنية على بيانات معروفة وذات صلة بالظاهرة.

الانحدار يفترض وجود علاقة خطية قوية.

تحليل الانحدار Regression Analysis:

الأساليب التي تستخدم في تقدير قيمة متغير عند معرفة قيم متغير آخر.

تحليل الانحدار الخطي الثنائي Bivariate Linear Regression: يستخدم لتمثيل العلاقة على شكل معادلة خطية للتنبؤ بقيمة متغير من خلال قيم متغير آخر، ويكون المتغير الأول كمياً ويسمى المتنبيء ويكون المتغير الثاني كمياً ويسمى المتنبأ به.

*** أهداف تحليل الانحدار Regression Analysis Objectives**

1. تحديد العلاقة بين المتغير التابع Y والمتغيرات المستقلة X_i
2. التنبؤ بقيمة المتغير التابع Y عن طريق المتغيرات المستقلة X_i
3. الاستنتاج حول المجتمع من خلال المعادلة التقديرية.
4. اختبار الفروق بين خط الانحدار التقديري وخط الانحدار الحقيقي.

*** الانحدار الخطي البسيط Simple Linear Regression**

هـــو الـــبحث في العلاقة بين متغيرين فقط وهما المتغير التابع Y والمتغير المستقل X، وإن معادلة الانحدار في المجتمع هي:

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Y: المتغير التابع. X: المتغير المستقل. α: معامل ثابت. β: ميل الانحدار.

ε_i: الأشياء الأخرى التي تسبب الخطأ Disturbance term error

المـــتغير المستقل Independent Variable: المتغير الذي يستخدم في تقدير قيم المتغير الآخر(السبب).

المتغير التابع Dependent Variable: المتغير الذي نقدره (الأثر).

التـــباين Variance: وجـــود فروق بين المشاهدات، مسافات، إذا كانت كبيرة يكون التباين كبير، وإذا كانت صغيرة يكون التباين صغير، وإذا انعدمت انعدم التباين.

إذا اعتمدت معادلة الانحدار على بيانات العينة نستعيض عن α بالرمز a وعن β بالرمز b وتصبح

معادلة الانحدار في العينة هي:

$$\hat{y}_i = a + bX_i + e_i$$

$$e = (Y - \overline{Y})$$

$$\hat{y} = a + bX_i$$

Y: المتغير التابع. X: المتغير المستقل. a: ثابت. b: معامل انحدار.

إن خط الانحدار يمر من خلال نقطة الوسط $(\overline{X}, \overline{Y})$، لذلك فإن مجموعة الأخطاء دائماً تساوي صفر $\sum e_i = 0$

*** معادلة خط الانحدار Regression Equation**

1- المعادلة التقديرية للانحدار الخطي البسيط Equation Estimation

$$\mathbf{Y_i = E(Y_i) + \varepsilon_i}$$
$$\mathbf{Y_i = a + bX_i + e_i}$$

ei : الخطأ العشوائي.

2- طريقة المربعات الصغرى Least Squares Method

حـــتى نقدر ميل خط الانحدار غير المعلوم بواسطة طريقة المربعات الصغرى والتي تعتمد على تقليل مجموع مربعات انحرافات القيم الحقيقية y_i عن القيم التقديرية $\hat{y}_i$ حيث (b) تمثل الميل

وهو نسبة تغير قيمة المتغير التابع y إلى وحدة واحدة من المتغير المستقل x وان a تمثل معامل الـــتقاطع والذي يعني مقدار قيمة y عندما تكون قيمة المتغير المستقل x = صفراً. وذلك بإتباع الخطوات التالية:

- معادلة خط الميل التقديري $\hat{y} = a + bX$ حيث $\hat{y}$, b,a, هي تقدير $\beta, \alpha, E(y)$
- انحراف كل قيمة حقيقية y_i عن القيمة التقديرية $\hat{y}_i$ يكون بمقدار e_i وهي $e_i = y_i - \hat{y}_i$

$$b = \frac{\sum X_i Y_i - n\bar{X}\,\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$a = \bar{Y} - b * \bar{X}$$

- ويمكـــن حساب معامل الانحدار b إذا علمنا بقيمة معامل الارتباط البسيط بين x, y كالتالي:

$$b = \frac{S_y}{S_x} * r_{xy}$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

- بتعويض قيم a, b نحصل على خط الانحدار التقديري كتقدير لخط انحدار المجتمع

$$E(Y) = \alpha + \beta X$$

أ- معادلة خط انحدار ص على س هي:

ص = أ + ب س	التنبؤ بقيمة ص إذا علمت قيمة س

ب = ($\sigma_{ص}$ / $\sigma_{س}$) * ر

$$ب = \frac{\sum(س_ر\, ص_ر) - ن * \bar{س} * \bar{ص}}{\sum س_ر^2 - ن * \bar{س}^2}$$

$$أ = \bar{ص} - ب\,\bar{س}$$

ب- معادلة خط انحدار س على ص هي:

س = أ + ب ص	التنبؤ بقيمة س إذا علمت قيمة ص

$$\text{ب} = (\sigma_{\text{س}} / \sigma_{\text{ص}}) * \text{ر}$$

$$\text{ب} = \frac{\sum \text{س}_{\text{ر}}\,\text{ص}_{\text{ر}} - \text{ن} * \bar{\text{س}} * \bar{\text{ص}}}{\sum \text{ص}_{\text{ر}}^{2} - \text{ن} * \bar{\text{ص}}^{2}}$$

$$\text{أ} = \bar{\text{س}} - \text{ب}\,\bar{\text{ص}}$$

* خواص مقدرات طريقة المربعات الصغرى

- خاصية عدم التحيز. خاصية أقل تباين ممكن.

* **الانحدار الخطي المتعدد** **Multiple Linear Regression**

الانحدار المتعدد (متغيرين مستقلين) Two Independent Variables

تحليل الانحدار ملائم لمعرفة كم من التباين يستطيع المتغيران معاً أن يفسرانه من التباين في المتغير التابع.

بمعرفة تباين المقدار الذي يشرحه المتغير الأول، ومعرفة تباين المقدار الذي يشرحه المتغير الثاني، نستطيع أن نكشف عن الأهمية النسبية لكلا المتغيرين.

الافتراضات للانحدار المتعدد

1. أن المتغيرات المستقلة مستقلة (معامل الارتباط بينها 0).
2. أن المتغيرات الداخلة على يمين إشارة=في المعادلة يمكن أن يكون بينها ترابطات

المتغيرات المستقلة (متنبئات) Predictors المتغـــير التابـــع (محك) Criterion

* **معادلة الانحدار الخطي المتعدد**

$$\hat{y} = a + b_1x_1 + b_2x_2$$

a : المقطع الصادي وهو بعد النقطة التي يقطع عندها خط الانحدار.

b_1: معامل الانحدار للمتغير الأول x_1 (وزن الانحدار)، تعني التغير في y الذي يقابل تغير x_1 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة.

b_2: معامل الانحدار للمتغير الثاني x_2، تعني التغير في y الذي يقابل تغير x_2 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة.

اختيار الأساليب الإحصائية الوصفية المستخدمة لمتغير واحد:

نوع المتغير	أساليب القياس المناسبة			
	نزعه مركزية	تشتت	المقاييس النسبية	أخرى
اسمي Nominal	المنوال	التكرار النسبي للقيمة المنوالية	التكرار النسبي (% للتكرارات)	------
رتبي Ordenal	الوسيط	نصف المدى الربيعي	التكرار النسبي	------
فئوي أو نسبي Interval Or Ratio	المتوسط للتوزيع الاعتدالي. الوسط والمتوسط للتوزيع ملتو	المدى المطلق. التباين الانحراف المعياري	التكرار النسبي مثل النسبة، الميئني الارباعيات	معاملات الالتواء والتفرطح.

5-1 استخدام برمجية SPSS في حل المسائل.

مثال 1-1)

اذا كانـــت الاســـتبانة التالية هي عبارة عن استجابة (10) طلاب على سؤال بتدريج ليكرت الخماسي وهو (ان التعليم العالي في الاردن في تطور مستمر)، وعلامات الطلاب بمادتي الاحصاء والقياس، كما تحتوي بعض المعلومات عن الطلبة المستجيبين لهذه الاستبانة.

o رقم الطالب : ....

الجنس ◊ ذكر ◊ انثى

المستوى ◊ اولى ◊ ثانية ◊ ثالثة ◊ رابعة

الرقم	الجنس	المستوى	علامة الاحصاء	علامة القياس	(ان التعليم العالي في الاردن في تطور مستمر) درجة الموافقة				
					موافق بشدة	موافق	محايد	معارض	معارض بشدة
1-	1	1	65	90		–			
2-	2	3	70	85		–			
3-	1	2	75	83		–			
4-	1	4	80	95	–				
5-	2	3	60	90			–		
6-	1	2	90	70					–
7-	1	3	80	75			–		
8-	2	2	65	60	–				
9-	2	2	80	80				–	
10-	1	1	55	50					–

اجب عن الاسئلة التالية:

1- اعمل على ترميز المتغيرات الواردة بالسؤال اعلاه.

2- ادخل البيانات الى البرنامج.

3- احسب مقاييس النزعة المركزية.

4- احسب مقاييس التشتت.

5- ما نسبة الذكور والإناث في عينة الدراسة؟

6- ما عدد أفراد العينة في كل مستوى؟

7- جد معامل الارتباط بين الجنس والمستوى وهل هو دال على مستوى ($\alpha=0.05$).

8- جد معامل الارتباط بين علامة الاحصاء وعلامة القياس وهل هو دال على مستوى ($\alpha=0.05$).

9- هل هناك علاقة بين مستوى الطالب ودرجة موافقته وهل هو دال على مستوى ($\alpha=0.05$) ولصالح من؟

10- اعمل على حفظ الملف تحت اسم Q1

الحل:

1- اعمل على ترميز المتغيرات الواردة بالسؤال اعلاه.

عملية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

- متغير الجنس Sex وهو (ذكر، أنثى)، حيث يعطى الرقم 1 للذكور والرقم 2 للإناث.

- المستوى الدراسي:

  - سنة اولى. 1 سنة ثانية. 2
  - سنة ثالثة. 3 سنة رابعة. 4

- أسئلة الاستبانة مصممة على أساس مقياس "ليكرت" الخماسي كما يلي:

  - موافق بشدة. 5
  - موافق. 4
  - محايد. 3
  - معارض. 2
  - معارض بشدة. 1

لعمل جدول ترميز للمتغيرات الواردة في مثال 1-1 ، بواسطة استخدام برنامج SPSS يتم كما يلي:

* تشغيل نظام **SPSS**

Start – Programs – SPSS for Windows – SPSS10.0for Windows – Type in data – Ok

start - Programs - SPSS for Windows - SPSS 10.0 for Windows - Type in data - OK

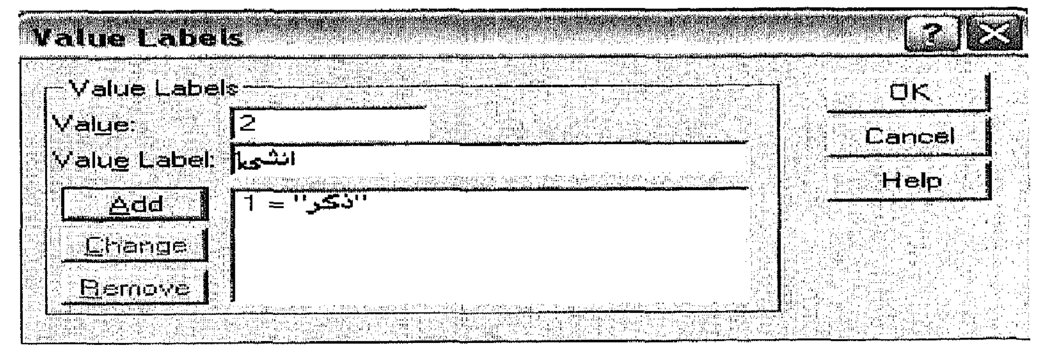
تظهر لديك الشاشة المبينة أدناه:

الخطوات المتبعه لتعريف المتغيرات: لنفرض أننا نريد تعريف المتغيرات الواردة في مثال1-1

انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه:

نعبيء الخانات أعلاه وهي Name ، Type ، ... ، Measure بالشكل التالي:

مثلا لادخال متغير الجنس نعبيء الخانات أعلاه وهي Name اسم المتغير ونكتب الجنس، ثم نوع المتغير Type ونختار Numeric أي رقمي، ... ، وعند وصولنا إلى خانة Valuesوهي القيم المحتملة للمتغير نعبأها بالشكل التالي:



وبعد الانتهاء من تعريف المتغيرات تظهر الشاشة كما هو مبين أدناه:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	الرقم	Numeric	8	0		None	None	8	Center	Scale
2	الجنس	Numeric	8	0		{1, ذكر}...	None	8	Center	Nominal
3	المستوى	Numeric	8	0		{1, سنة اولى}...	None	8	Center	Ordinal
4	سن	Numeric	8	0		{1, [illegible]}	None	8	Center	Ordinal
5	احصاء	Numeric	8	0		None	None	8	Center	Scale
6	فيلبس	Numeric	8	0		None	None	8	Center	Scale

انقر على Data View الموجودة على شريط الحالة فتظهر الشاشة أدناه:

مثال 1-1 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

10 : الجنس

	الرقم	الجنس	المستوى	سن1	احصاء	فيلبس	[illegible]
1							
2							
3							
4							
5							

2- ادخل البيانات الى البرنامج.

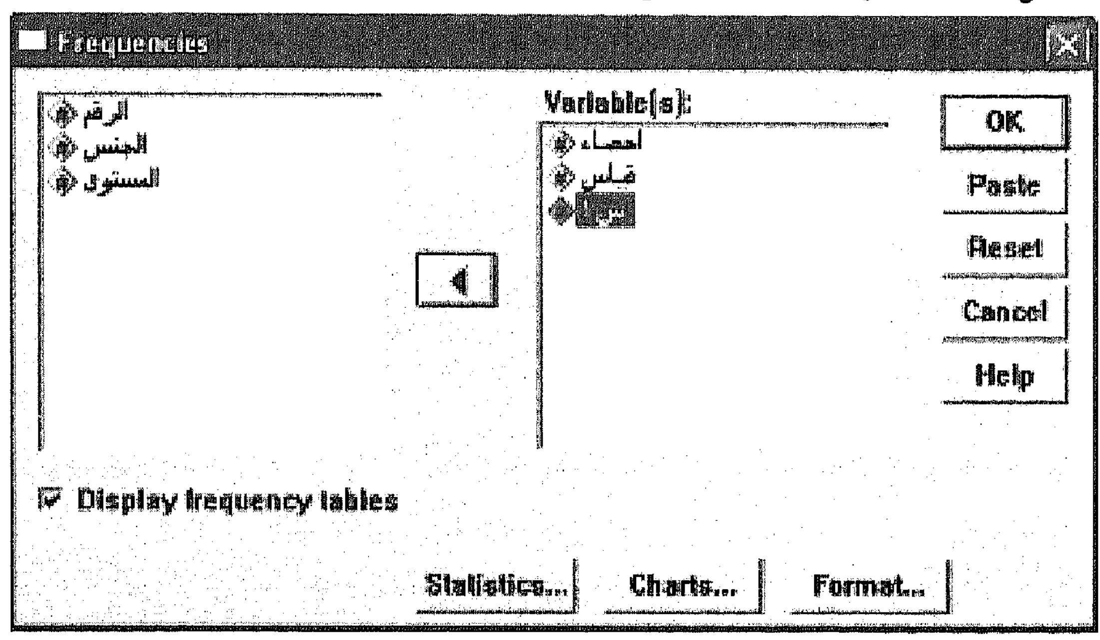
إدخال البيانات Input Data

	الرقم	الجنس	المستوى	سن1	احصاء	فيلبس
1	1	1	1	4	65	90
2	2	2	3	4	70	85
3	3	1	2	4	75	83
4	4	1	4	5	80	95
5	5	2	3	3	60	90
6	6	1	2	1	90	70
7	7	1	3	3	80	75
8	8	2	2	5	65	60
9	9	2	2	2	80	80
10	10	1	1	1	55	50

بعد الانتهاء من ادخال البيانات.

3- احسب مقاييس النزعة المركزية.

لعمل إحصاءات وصفية...Analyze – Descriptive Statistics – Frequencies



نعمـــل علـــى وضع المتغيرات احصاء، قياس، س1 في قائمة :Variable(s) ثم نضغط زر Statistics فتظهر الشاشة أدناه:

Frequencies: Statistics

Percentile Values
Quartiles
Cut points for 10 equal groups
Percentile(s):
Add
Change
Remove

Central Tendency
Mean
Median
Mode
Sum

Continue
Cancel
Help

Values are group midpoints

Dispersion
Std. deviation
Variance
Range
Minimum
Maximum
S.E. mean

Distribution
Skewness
Kurtosis

نحدد مقاييس النزعة المركزية Central Tendency ثم نضغط زر Continue

نضغط زر Charts فتظهر الشاشة أدناه:

Frequencies: Charts

Chart Type
- None
- Bar charts
- Pie charts
- Histograms
  - With normal curve

Continue
Cancel
Help

Chart Values
- Frequencies
- Percentages

نحدد Histograms ونحدد with normal curve ثم نضغط زر Continue

ثم نضغط زر ok فتظهر شاشة المخرجات أدناه:

Statistics

		احصاء	قياس	س١
N	Valid	10	10	10
	Missing	0	0	0
Mean		72.00	77.80	3.20
Median		72.50	81.50	3.50
Mode		80	90	4
Sum		720	778	32

Frequency Table

احصاء

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	55	1	10.0	10.0	10.0
	60	1	10.0	10.0	20.0
	65	2	20.0	20.0	40.0
	70	1	10.0	10.0	50.0
	75	1	10.0	10.0	60.0
	80	3	30.0	30.0	90.0
	90	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

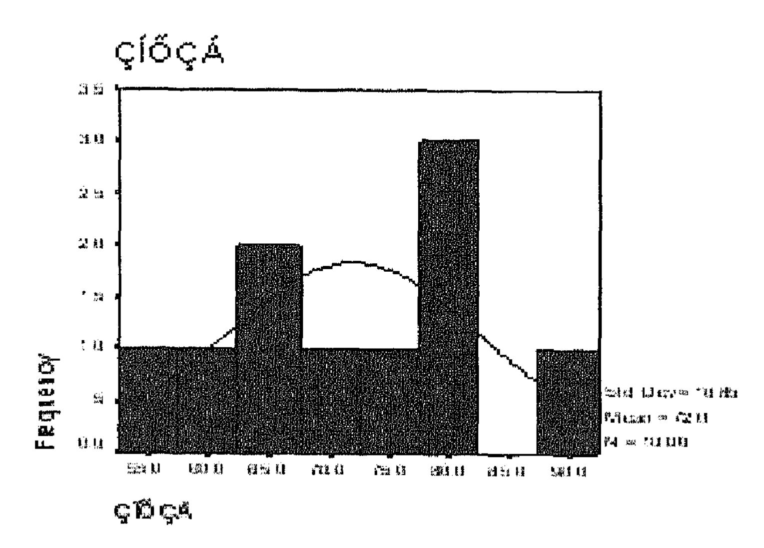
قياس

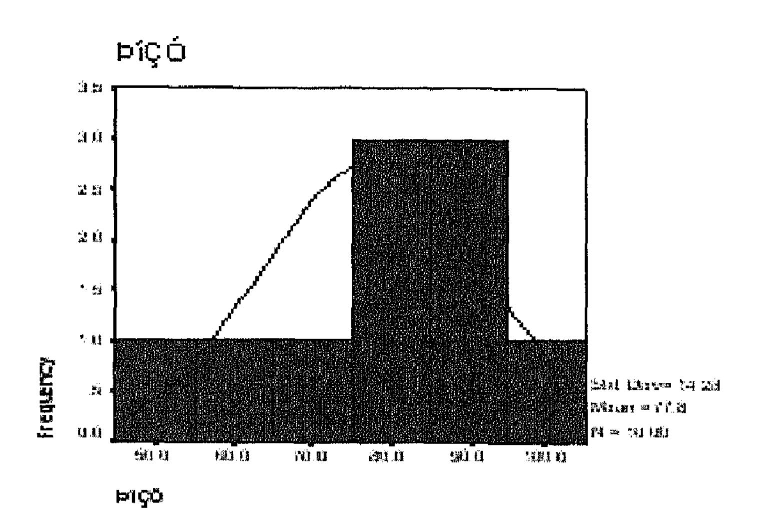
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	50	1	10.0	10.0	10.0
	60	1	10.0	10.0	20.0
	70	1	10.0	10.0	30.0
	75	1	10.0	10.0	40.0
	80	1	10.0	10.0	50.0
	83	1	10.0	10.0	60.0
	85	1	10.0	10.0	70.0
	90	2	20.0	20.0	90.0
	95	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

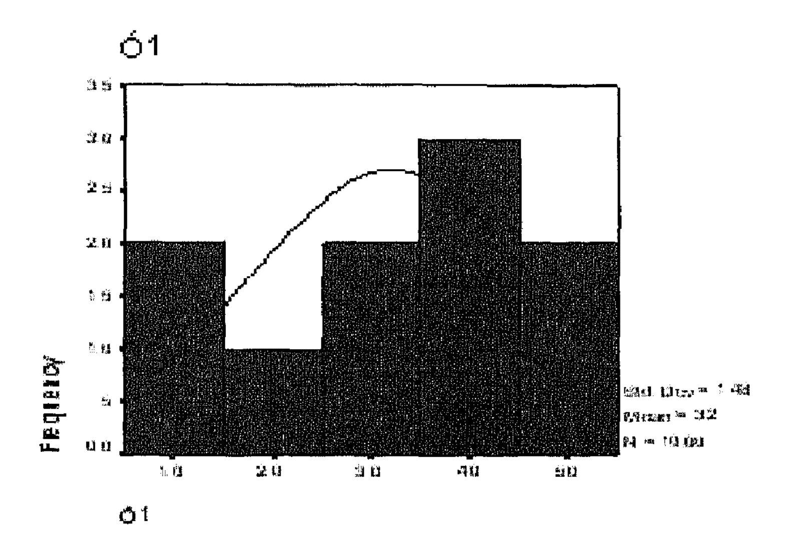
س١

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	معارض بشدة	2	20.0	20.0	20.0
	معارض	1	10.0	10.0	30.0
	محايد	2	20.0	20.0	50.0
	موافق	3	30.0	30.0	80.0
	موافق بشدة	2	20.0	20.0	100.0
	Total	10	100.0	100.0	

الرسوم البيانية للمتغيرات







4- احسب مقاييس التشتت.

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies

نعمـــل علــــى وضع المتغيرات احصاء، قياس، س1 في قائمة :Variable(s) ثم نضغط زر Statistics ونحدد مقاييس التشتت Dispersionثم نضغط زرContinue ثم OK

Statistics

		احصاء	قياس	س1
N	Valid	10	10	10
	Missing	0	0	0
Std. Error of Mean		3.432	4.516	.467
Std. Deviation		10.853	14.281	1.476
Variance		117.778	203.956	2.178
Range		35	45	4
Minimum		55	50	1
Maximum		90	95	5

احصاء

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	55	1	10.0	10.0	10.0
	60	1	10.0	10.0	20.0
	65	2	20.0	20.0	40.0
	70	1	10.0	10.0	50.0
	75	1	10.0	10.0	60.0
	80	3	30.0	30.0	90.0
	90	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

قياس

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	50	1	10.0	10.0	10.0
	60	1	10.0	10.0	20.0
	70	1	10.0	10.0	30.0
	75	1	10.0	10.0	40.0
	80	1	10.0	10.0	50.0
	83	1	10.0	10.0	60.0
	85	1	10.0	10.0	70.0
	90	2	20.0	20.0	90.0
	95	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

س1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	معارض بشدة	2	20.0	20.0	20.0
	معارض	1	10.0	10.0	30.0
	محايد	2	20.0	20.0	50.0
	موافق	3	30.0	30.0	80.0
	موافق بشدة	2	20.0	20.0	100.0
	Total	10	100.0	100.0	

5- ما نسبة الذكور والإناث في عينة الدراسة؟

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies

نعمل على وضع المتغير (الجنس) في قائمة :Variable(s) ثم نضغط ثم OK

الجنس

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ذكر	6	60.0	60.0	60.0
	انثى	4	40.0	40.0	100.0
	Total	10	100.0	100.0	

6- ما عدد أفراد العينة في كل مستوى؟

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies

نعمل على وضع المتغير (المستوى) في قائمة :Variable(s) ثم نضغط ثم OK

المستوى

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	سنةأولى	2	20.0	20.0	20.0
	سنةثانية	4	40.0	40.0	60.0
	سنةثالثة	3	30.0	30.0	90.0
	سنةرابعة	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

7- جد معامل الارتباط بين الجنس والمستوى وهل هو دال على مستوى ($\alpha=0.05$).

لعمل إحصاءات وصفية Analyze – Descriptive Statistics – Crosstabs...

Crosstabs

الرقم
س 1
اسماء
جنس

Row(s): الجنس

Column(s): المستوى

Layer 1 of 1

OK | Paste | Reset | Cancel | Help

Display clustered bar charts

Suppress tables

Statistics... | Cells... | Format...

ثم نضغط ثم OK

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
الجنس * المستوى	10	100.0%	0	.0%	10	100.0%

الجنس * المستوى Crosstabulation

Count

		المستوى				Total
		[illegible]	[illegible]	[illegible]	[illegible]	
الجنس	ذكر	2	2	1	1	6
	انثى		2	2		4
Total		2	4	3	1	10

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.553	.383
	Cramer's V	.553	.383
N of Valid Cases		10	

a. Not assuming the null hypothesis.

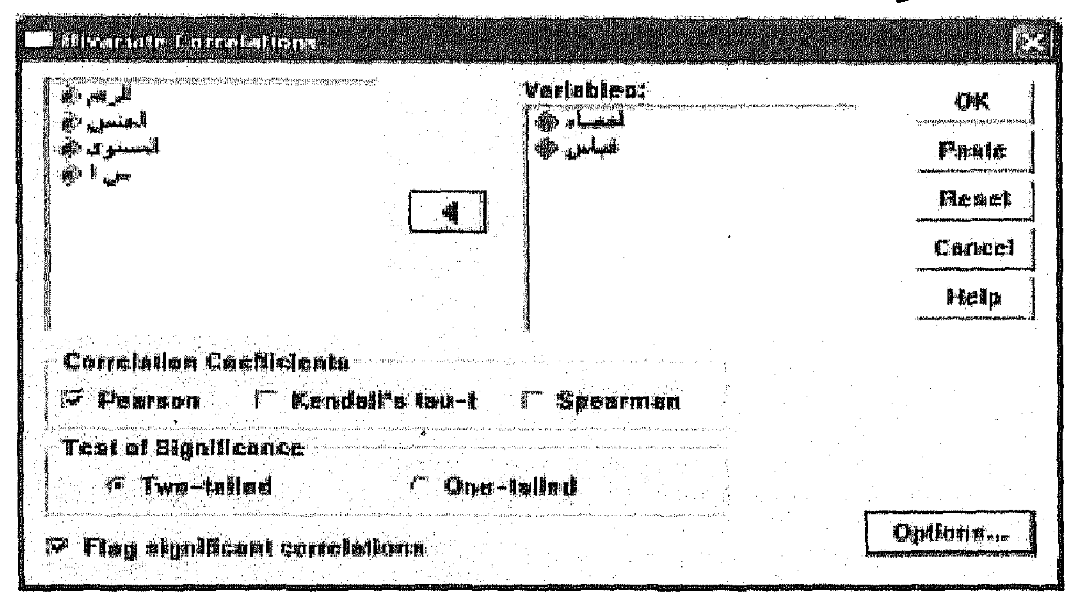
b. Using the asymptotic standard error assuming the null hypothesis.

معامل الارتباط فاي بين الجنس والمستوى = 0.553

وهو غير دال احصائياً لأن Sig=0.383 وهي اكبر من مستوى الدلالة ($\alpha=0.05$)

8- جد معامل الارتباط بين علامة الاحصاء وعلامة القياس وهل هو دال على مستوى (α=0.05).

لحساب معامل الارتباط Analyze – Correlate – Bivariate...



ثم نضغط ثم OK

➔ Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
احصاء	72.00	10.853	10
قياس	77.80	14.281	10

Correlations

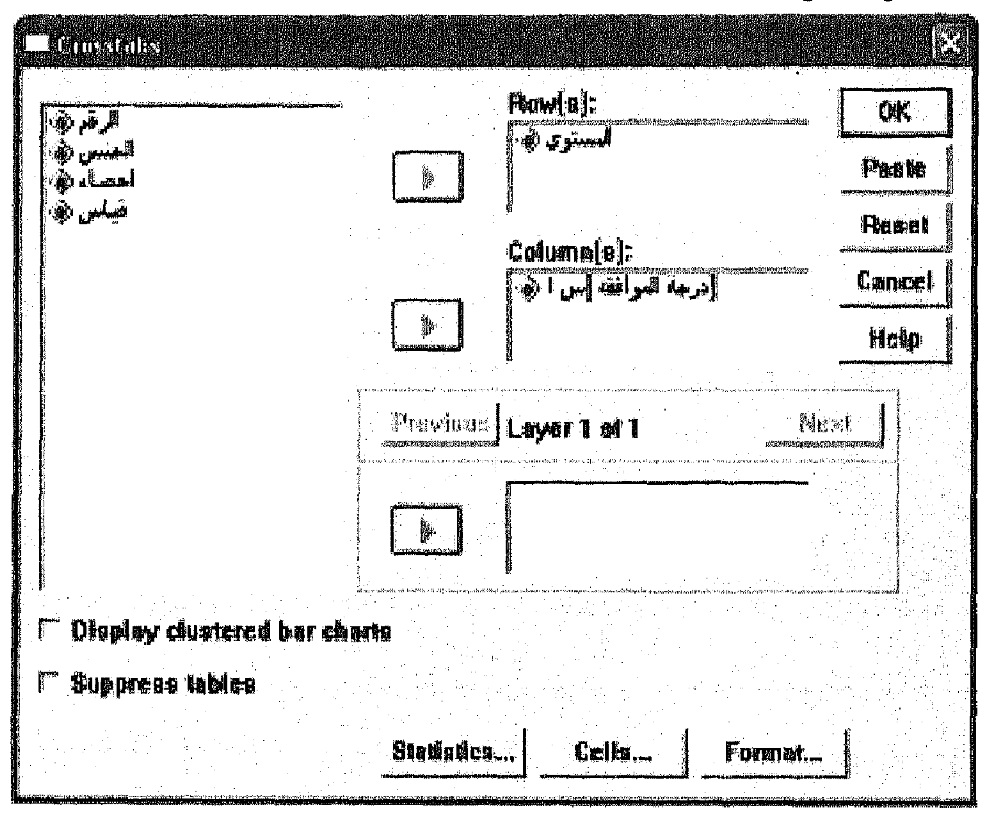
		احصاء	قياس
احصاء	Pearson Correlation	1	.257
	Sig. (2-tailed)	.	.473
	N	10	10
قياس	Pearson Correlation	.257	1
	Sig. (2-tailed)	.473	.
	N	10	10

معامل الارتباط بيرسون بين الاحصاء والقياس = 0.257

وهو غير دال احصائياً لأن Sig=0.473 وهي اكبر من مستوى الدلالة (α=0.05)

9- هــل هــناك علاقــة بين مستوى الطالب ودرجة موافقته وهل هو دال على مستوى ($\alpha=0.05$) ولصالح من؟

لعمل معامل ارتباط ...Analyze – Descriptive Statistics – Crosstabs



نضغط زر ...Statistics

Crosstabs: Statistics

Chi-square
Correlations
Continue
Cancel
Help
Nominal
Contingency coefficient
Phi and Cramér's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals: 1

❖ Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
المستوى * درجة الموافقة	10	100.0%	0	.0%	10	100.0%

المستوى * درجة الموافقة Crosstabulation

Count

		درجة الموافقة					Total
		معارض بشدة	معارض	محايد	موافق	موافق بشدة	
المستوى	سنة اولى	1			1		2
	سنة ثانية	1	1		1	1	4
	سنة ثالثة			2	1		3
	سنة رابعة					1	1
Total		2	1	2	3	2	10

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Gamma	.400	.303	1.222	.222
N of Valid Cases		10			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Correlations

			المستوى	درجة الموافقة
Spearman's rho	المستوى	Correlation Coefficient	1.000	.361
		Sig. (2-tailed)	.	.306
		N	10	10
	درجة الموافقة	Correlation Coefficient	.361	1.000
		Sig. (2-tailed)	.306	.
		N	10	10

معامل الارتباط جاما بين مستوى الطالب ودرجة الموافقه = 0.400

وهو غير دال احصائياً لأن Sig=0.222 وهي اكبر من مستوى الدلالة ($\alpha=0.05$)

معامل الارتباط سبيرمان بين مستوى الطالب ودرجة الموافقه = 0.361

وهو غير دال احصائياً لأن Sig=0.306 وهي اكبر من مستوى الدلالة ($\alpha=0.05$)

10- اعمل على حفظ الملف تحت اسم مثال1-1

لحفظ الملف من قائمة File – Save As....

: Save Data As

Save in: كتاب الاحصاء التحليلي

استخدام الحاسوب-الفصل الأول-٢٠٠٩

fa

saq

الفصل الاول

مثال١-١

Keeping 6 of 6 variables.

Variables...

File name: مثال1-1

Save

Save as type: مثال١-١.sav

Paste

Write variable names to spreadsheet

Cancel

ثم نضغط زر Save

مثال1-2:

في اختـــبار لمـــادة الحاسوب الذي يتكون من (15) سؤال من نوع الاختيار من متعدد والذي اجاب عليه (20) طالب، والمطلوب منك تحليل هذا الاختبار والتعليق على نتائجه؟

رقم الطالب: التخصص: 1-علمي 2-أدبي 3-معلوماتية 4- مهني

الجنـــس : 1-ذكر 2-أنثى .

المادة: الحاسوب. المستوى الثالث: تراسل البيانات والشبكات .

	stno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1

* أوجد ما يلي: مع إعطاء التعليق المناسب.

1- اعمل على ترميز المتغيرات الواردة في السؤال Coding

2- اعمل على ادخال البيانات الخاصة بالاختبار كما هو مبين ادناه

3- علامة كل طالب على الاختبار.

4- الوسط الحسابي والانحراف المعياري لكل سؤال في الاختبار.

5- الوسط الحسابي والانحراف المعياري للاختبار.

6- معامل الثبات للاختبار.

7- معامل الارتباط بين السؤال الاول والمجموع وهل هو دال على مستوى ($\alpha=0.01$)

8- اعمل على حفظ الملف تحت اسم مثال1-2

1– اعمل على ترميز المتغيرات الواردة في السؤال Coding

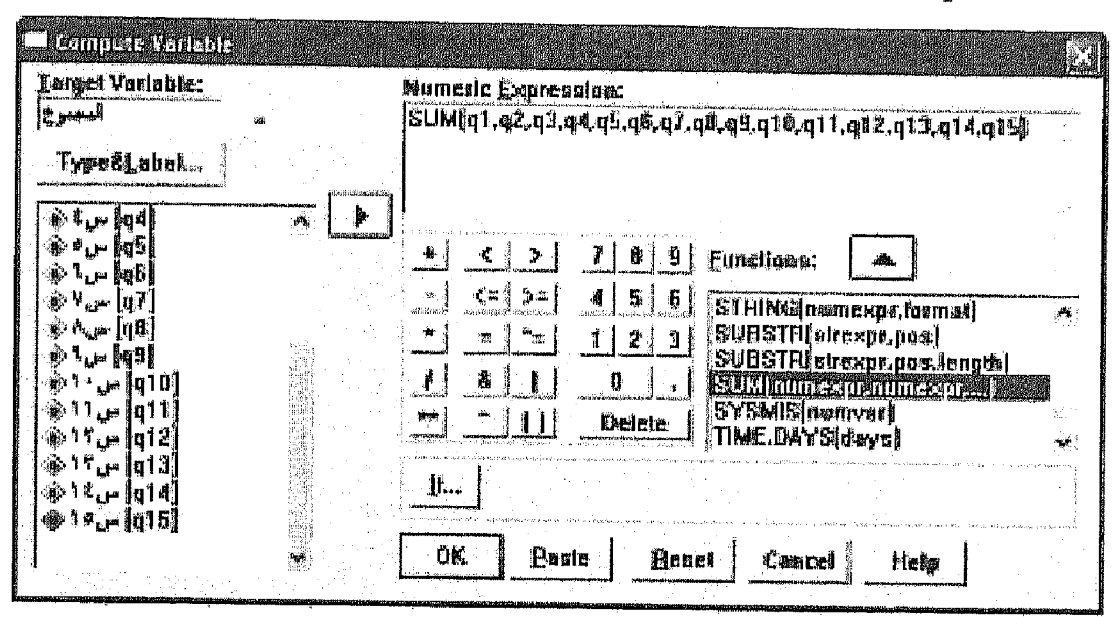
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	sino	Numeric	3	0	رقم الطالب	None	None	4	Center	Scale
2	sex	Numeric	1	0	الجنس	{1, ذكر}...	None	3	Center	Nominal
3	spec	Numeric	1	0	التخصص	{1, علمي}...	None	3	Center	Nominal
4	q1	Numeric	1	0	س١	None	None	3	Center	Scale
5	q2	Numeric	1	0	س٢	None	None	3	Center	Scale
6	q3	Numeric	1	0	س٣	None	None	4	Center	Scale
7	q4	Numeric	1	0	س٤	None	None	4	Center	Scale
8	q5	Numeric	1	0	س٥	None	None	4	Center	Scale
9	q6	Numeric	1	0	س٦	None	None	4	Center	Scale
10	q7	Numeric	1	0	س٧	None	None	4	Center	Scale
11	q8	Numeric	1	0	س٨	None	None	4	Center	Scale
12	q9	Numeric	1	0	س٩	None	None	4	Center	Scale
13	q10	Numeric	1	0	س١٠	None	None	4	Center	Scale
14	q11	Numeric	1	0	س١١	None	None	3	Center	Scale
15	q12	Numeric	1	0	س١٢	None	None	4	Center	Scale
16	q13	Numeric	1	0	س١٣	None	None	3	Center	Scale
17	q14	Numeric	1	0	س١٤	None	None	4	Center	Scale
18	q15	Numeric	1	0	س١٥	None	None	4	Center	Scale

2– اعمل على ادخال البيانات الخاصة بالاختبار كما هو مبين ادناه

	sino	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1

3- علامة كل طالب على الاختبار.

Transform – Compute...

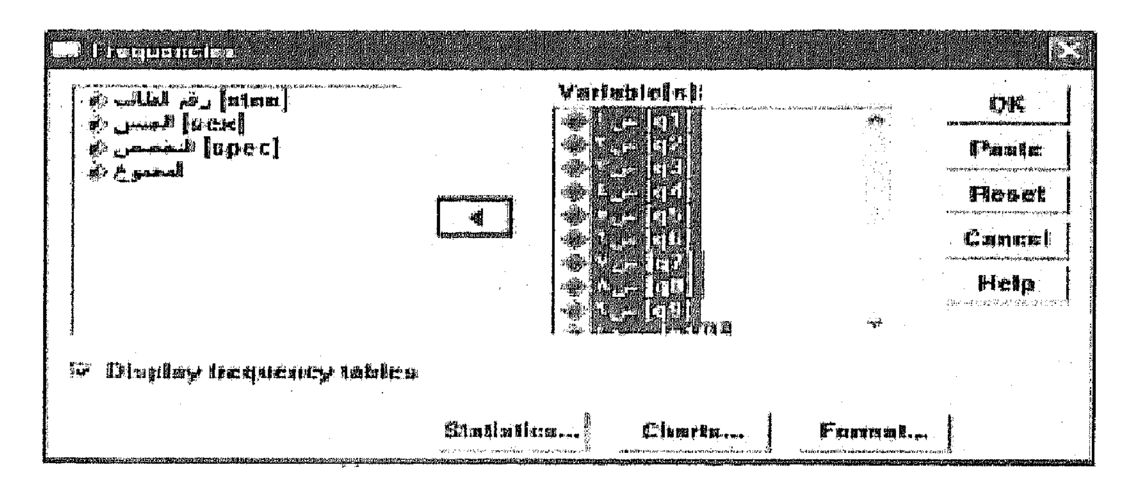


ثم نضغط Ok

	stno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15	المجموع
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	13.00
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	14.00
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	13.00
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	13.00
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	12.00
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	12.00
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	10.00
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1	10.00
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1	11.00
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	12.00
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	14.00
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	14.00
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	13.00
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	6.00
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	10.00
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	12.00
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	12.00
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0	8.00
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	14.00
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1	9.00

4- الوسط الحسابي والانحراف المعياري لكل سؤال في الاختبار.

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies



نضغط زر ...Statistics

Frequencies: Statistics

Percentile Values
Quartiles
Cut points for 10 equal groups
Percentile(s):
Add
Change
Remove

Central Tendency
Mean
Median
Mode
Sum

Continue
Cancel
Help

Values are group midpoints

Dispersion
Std. deviation
Variance
Range
Minimum
Maximum
S.E. mean

Distribution
Skewness
Kurtosis

نضغط زر ...Continue ، ثم نضغط زر Ok

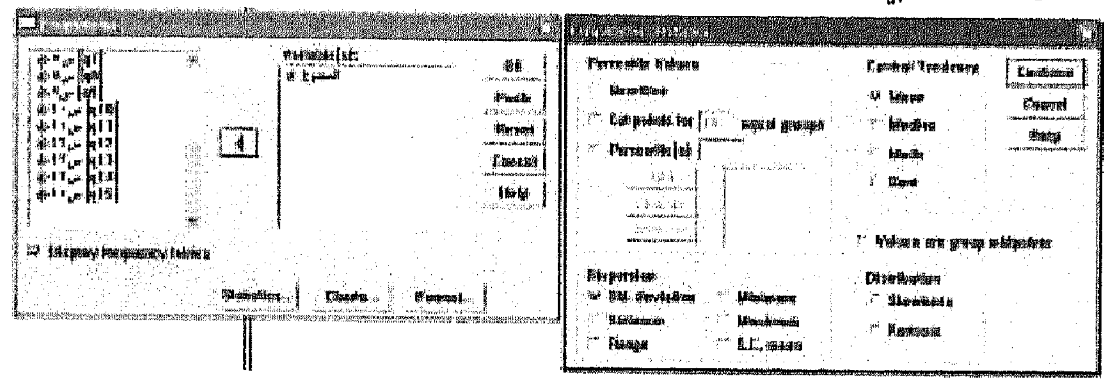
Statistics

		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
N	Valid	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Missing	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Mean		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Std. Deviation		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Statistics

		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
N	Valid	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Missing	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Mean		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Std. Deviation		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

5- الوسط الحسابي والانحراف المعياري للاختبار.



نضغط زر ...Continue ، ثم نضغط زر Ok

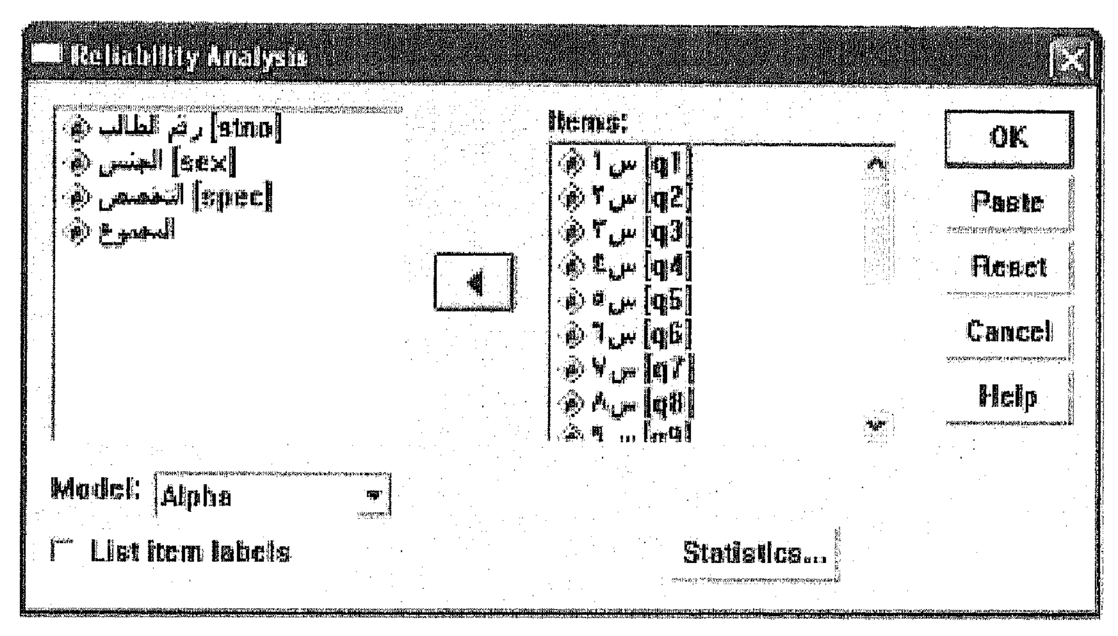
Statistics

المجموع

N	Valid	20
	Missing	0
Mean		11.6000
Std. Deviation		2.18608

6- معامل الثبات للاختبار.

Analyze – Scale – Reliability Analysis



ثم نضغط زر Ok

→ **Reliability**

```
****** Method 1 (space saver) will be used for this analysis ******

R E L I A B I L I T Y   A N A L Y S I S  -  S C A L E   (A L P H A)

                          Mean        Std Dev       Cases

  1.     Q1              .8500         .3663         20.0
  2.     Q2              .8000         .4104         20.0
  3.     Q3              .1500         .3663         20.0
  4.     Q4              .8000         .4104         20.0
  5.     Q5              .8000         .4104         20.0
  6.     Q6              .8000         .4104         20.0
  7.     Q7              .9000         .3078         20.0
  8.     Q8             1.0000         .0000         20.0
  9.     Q9             1.0000         .0000         20.0
 10.     Q10             .8500         .3663         20.0
 11.     Q11             .8000         .4104         20.0
 12.     Q12             .7000         .4702         20.0
 13.     Q13             .9000         .3078         20.0
 14.     Q14            1.0000         .0000         20.0
 15.     Q15             .9500         .2236         20.0

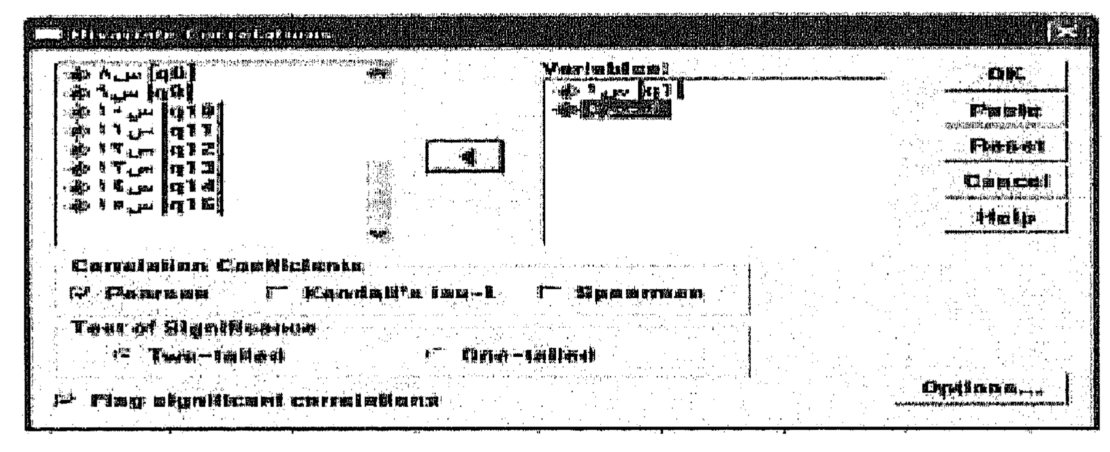
Reliability Coefficients

N of Cases =      20.0                    N of Items = 15

Alpha =     .6478
```

7- معامل الارتباط بين السؤال الاول والمجموع وهل هو دال على مستوى (α=0.01).

لحساب معامل الارتباط   Analyze – Correlate – Bivariate…



ثم نضغط زر Ok

Correlations

		س١	المجموع
س١	Pearson Correlation	1	.710**
	Sig. (2-tailed)	.	.000
	N	20	20
المجموع	Pearson Correlation	.710**	1
	Sig. (2-tailed)	.000	.
	N	20	20

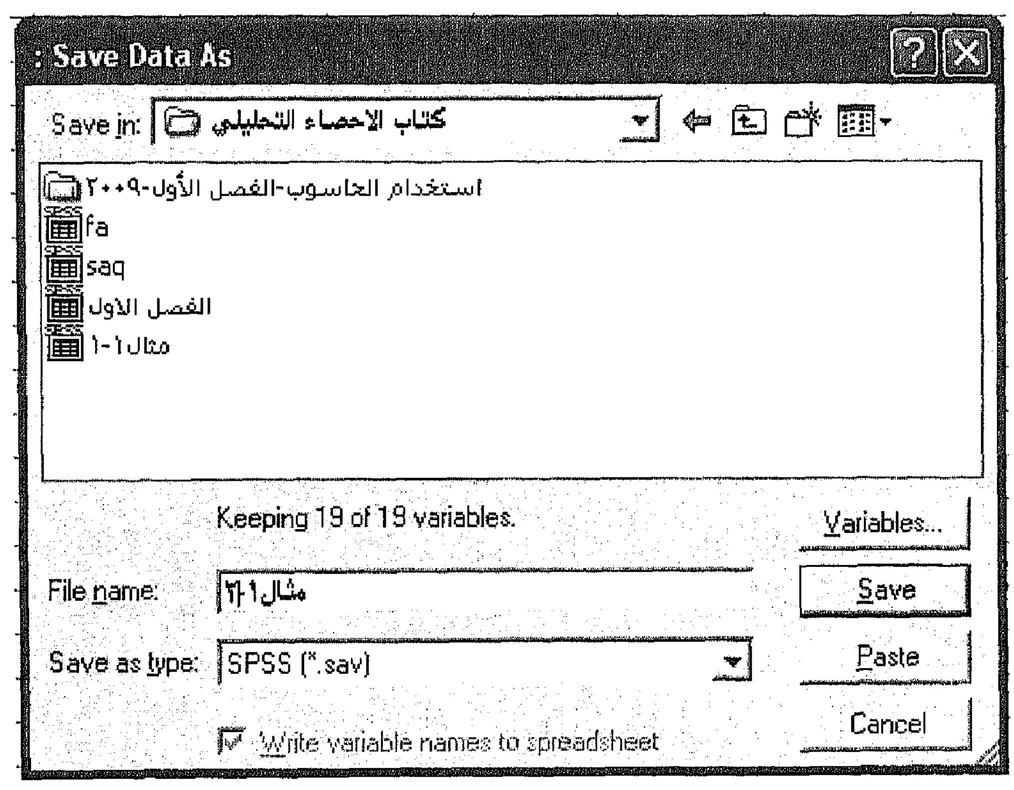
**. Correlation is significant at the 0.01 level

معامل الارتباط بيرسون بين السؤال الاول والمجموع = .7100

وهو ذو دلالة احصائية لأن Sig=0.000 وهي أقل من مستوى الدلالة (α=0.01)

8- اعمل على حفظ الملف تحت اسم مثال1-2

لحفظ الملف من قائمة File – Save As….



ثم نضغط زر Save

6-1 تمـاريـــن Exercise.

س1: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										
الرقم	21	22	23	24	25	26	27	28	29	30
الإجابة										

1- البيانات المبوبة هي البيانات:

أ- الأولية ب- المجدولة ج- الرقمية د- غيرالرقمية

2- البيانات النوعية هي التي تصف الظاهرة بشكل:

أ- أولي ب- نهائي ج- غير رقمي د- رقمي.

3- إن علامة الطالب هي مثال على البيانات:

أ- المبوبة ب- رمزي ج- النوعية د- الكمية.

4- إن جنس الطالب هو مثال على البيانات:

أ- المنفصلة ب- المتصلة ج- النوعية د- الكمية.

5- مؤشر المجتمع يسمى:

أ- احصائي ب- معلم ج- مؤشر مجتمع د- مؤشر.

6- مؤشر العينة يسمى:

أ- مؤشر ب- معلم عينة ج- معلم د- إحصائي.

7- إن الرمز التالي S_x معناه:

أ-تباين مجتمع ب- انحراف معياري مجتمع

ج-تباين عينة د- انحراف معياري عينة

8- تصنف المتغيرات حسب طبيعة المعلومات التي يؤديها القياس إلى:

أ- اسمية، ترتيبية، فئوية، نسبية.   ب- نوعية، كمية.

ج- مستقلة، تابعة.   د- رقمية، رمزية.

9- تسمى المتغيرات التي لها عدد فئات محدد من دون أي وزن لهذه الفئات ولا يوجد أفضلية لأحدها على الآخر بالمتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

10- إن متغير الجنس الذي يصنف المجتمع إلى فئتين هما الذكور والإناث هو مثال على المتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

11- الأرقام التي ليس لها معنى حقيقي ولا يمكن إجراء العمليات الحسابية عليها هي مثال على المتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

12- تسمى المتغيرات التي لها عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً، ولا يمكن تحديد الفروق بدقة بين القيم المختلفة بالمتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

13- إذا كانت A أكبر من B ولكن لا نستطيع معرفة كم يكبر A عن B هي مثال على المتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

14- المتغيرات الكمية التي يمكن إجراء العمليات الحسابية عليها وذلك دون أن تتأثر المسافة النسبية بين قيمها، ويميز هذا المتغير من خلال قيمة الصفر التي لا تعني عدم توافر تلك الصفة تسمى بالمتغيرات.

أ- اسمية   ب- ترتيبية   ج- فئوية   د- نسبية.

15- إذا كانت علامة جمال في مادة اللغة العربية أكثر من علامة أيمن، وأن علامة أيمن أكثر من علامة محمد، فإننا نعرف هذه المتغيرات بالمتغيرات.

أ- نسبية ب- فئوية ج- ترتيبية د- اسمية

16- المتغيرات الكمية التي ليس لها فئات محددة ولكن الصفر فيها يمثل عدم توفر الصفة، مثل المتغيرات الزمنية تسمى بالمتغيرات.

أ- اسمية ب- ترتيبية ج- فئوية د- نسبية.

17- إذا أراد مدرس أن يبحث عن أثر عدد ساعات الدراسة على تحصيل الطالب في مبحث معين، إن المتغير المستقل هو؟

أ- المدرس ب- ساعات الدراسة ج- التحصيل د- الطالب.

18- إذا أراد مدرس أن يبحث عن أثر عدد ساعات الدراسة على تحصيل الطالب في مبحث معين، إن المتغير التابع هو؟

أ- المدرس ب- ساعات الدراسة ج- التحصيل د- الطالب.

19- العينة التي يتم اختيارها بعناية وبصورة غير عشوائية تسمى بالعينة.

أ- القصدية ب- الطبقية ج- العنقودية د- المنتظمة.

20- العينة التي يتم اختيار أفرادها من المجتمع بحيث يكون لأي فرد من الأفراد الفرصة نفسها للظهور في هذه العينة تسمى بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المنتظمة.

21- إذا كان حجم المجتمع كبير جداً، وقسم المجتمع إلى مجموعات صغيرة، ثم اختيرت عينة عشوائية من هذه المجموعات الصغيرة تسمى العينة بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المنتظمة.

22- العينة التي تمثل المجتمع الاحصائي تمثيلاً صادقاً وتتفق مقاييسها الإحصائية مع مقاييس المجتمع ويتم اختيارها بصورة تتابعية تسمى بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المعيارية.

23- إذا كان السبب في أخطاء العينات هو خطأ في طريقة اختيار العينة، تسمى أخطاء العينة في هذه الحالة بأخطاء:

أ- عشوائية ب- التحيز ج- العينة د- المعطيات.

24- إذا كان السبب في أخطاء العينات هو زيادة أو نقص في البيانات، تسمى أخطاء العينة في هذه الحالة بأخطاء:

أ- عشوائية ب- التحيز ج- العينة د- المعطيات.

25- نتائج العينات تتعرض لمجموعة من الأخطاء منها ما يلي:

أ- المنتظمة ب- التحيز ج- العينة د-جميع ما ذكر

26- في تجـــربة لمعرفة أثر تخصص الطالب على تحصيله في مساق مهارات الحاسوب، ينظر إلى متغير تخصص الطالب كمتغير:

أ- مستقل ب- تابع ج- نسبي د- فئوي

27- رقم الشعبة في مساق متعدد الشعب مثال على المتغير:

أ- رتبي ب- الاسمي ج- نسبي د- فئوي

28- المتغير الذي لمقياسه صفر مطلق هو المتغير:

أ- رتبي ب- الاسمي ج- نسبي د- فئوي

29- التعريف التالي:"جزء من المجتمع تتم دراسة الظاهرة من خلال المعلومات عنها".

أ- الإحصاء ب- المجتمع ج- العينة د- القياس

30- يختلف الإحصاء الوصفي عن الإحصاء التحليلي بشكل أساسي في:

أ- الطريقة ب- الدقة ج-التعميم د-نوع البيانات

س2: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										
الرقم	21	22	23	24	25	26				
الإجابة										

1- الوسط الحسابي للقيم 0، 5، 1، -2، 6 هو:-

أ) 0 ب) -2 ج) 2 د) 5

2- ما قيمة الوسيط لمجموعة القيم التالية: 16، 18، 20، 24، 32، 22

أ) 11 ب) 16 ج) 20 د) 21

3- إذا كان وسيط (15) علامة هو (60) وعدلت العلامات حسب المعادلة ص=0.9س+3

حيث س العلامة السابقة، ص العلامة الجديدة، فإن الوسيط الجديد

أ) 54 ب) 15 ج) 57 د) 17.5

4- يعرف المئين 80 بأنه:

أ) قيمة في التوزيع = 80

ب) قيمة في التوزيع أكبر من 80% من مجموع القيم

ج) قيمة في التوزيع تزيد عن 80% من القيم

د) قيمة في التوزيع فوق الوسيط بمقدار 30 علامة

5- أي من المقاييس التالية يتأثر بالقيم المتطرفة :

أ) الوسط ب) الوسيط ج) المنوال د) لا شيء مما ذكر

6- أي من المقاييس التالية يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة):

أ) الوسط ب) الوسيط ج) المنوال د) جميع ما ذكر

7- أي من المقاييس التالية لا يعتمد على قيم البيانات إنما يعتمد على ترتيبها وموقعها:

أ) الوسط ب) الوسيط ج) المنوال د) لا شيء مما ذكر

استعن بالجدول للإجابة عن الأسئلة التالية:- (8، 9، 10، 11، 12)

التكرار التراكمي	الحدود الفعلية	المراكز*التكرار	المراكز	تكرار	الفئات
1	-0.5 – 4.5	2	2	1	00-04
3	4.5 – 9.0	14	7	2	05-09
6	9.5 – 14.5	36	12	3	10-14
8	14.5 – 19.5	34	17	2	15-19
10	19.5 – 24.5	44	22	2	20-24
		130		10	المجموع

8- مركز الفئة الرابعة هو:-

أ) 12 ب) 17 ج) 22 د) لا يمكن معرفته

9- الوسط الحسابي للتوزيع ($\overline{X}$)=

أ) 6 ب) 10 ج) 13 د) 60

10- المئين 60 للتوزيع =

أ) 14.5 ب) 12.5 ج) 9.5 د) 3

11- الحدود الفعلية للفئة الخامسة هي :

أ) 24-20 ب) 24.5-19.5

ج) 23.5-20.5 د) 24.5-20.5

12- طول الفئة للفئة الرابعة يساوي :

أ) 5 ب) 4 ج) 3 د) لاشيء مما ذكر

13- ما المنوال للقيم التالية: 4، 7، 2، 9، 2، 11

أ) 2 ب) 7 ج) 9 د) 11

14- إن مقاييس النزعة المركزية تتأثر بالعمليات التالية:

أ) الجمع ب) الطرح ج) الضرب د) جميع ما ذكر

15- منوال التوزيع التالي 3، 2، 2، 6 ، 7 ، 7 هو .....

أ) 2 ب) 7 ج) 3 د) 2، 7

16- أكثر مقاييس النزعة المركزية تأثراً بالقيم المتطرفة هو .....

أ) الوسط ب) الوسيط ج) المنوال د) المئين

17- أفضل مقاييس النزعة المركزية لوصف المتغيرات الأسمية هو ....

أ) الوسط ب) الوسيط ج) المنوال د) المئين

18- أقل مقاييس النزعة المركزية تأثراً بتقلبات العينة هو ....

أ) الوسط ب) الوسيط ج) المنوال د) المئين

19- أي مقاييس النزعة المركزية يقسم التوزيع إلى قسمين متساويين مهما كان شكله

أ) الوسط ب) الوسيط ج) المنوال د) المئين

20- إذا علمت أن الوسط الحسابي لعلامات طلاب صف ما هي 10 والانحراف المعياري 2 فإذا أضيفت 3 علامات لكل طالب فإن الوسط الحسابي للعلامات بعد الإضافة هو ....

أ) 10 ب) 2 ج) 3 د) 13

21- إن الوسط الحسابي لخمس قيم ثلاث منها متوسطها 4 والباقي متوسطها 5 هو:

أ) 10 ب) 12 ج) 24 د) 4.4

22- احسب الوسط الحسابي للقيم التالية : 12، 15، 14، 15، 22

أ) 78 ب) 15.6 ج) 15 د) 3

23- احسب الوسيط للقيم التالية : 12، 15، 14، 15، 22

أ) 78 ب) 15.6 ج) 15 د) 3

24- احسب المنوال للقيم التالية : 12، 15، 14، 15، 22

أ) 78 ب) 15.6 ج) 15 د) 3

25- إذا علمت أن الوسط الحسابي لعلامات طلاب مساق الإحصاء هو 20 بينما كان الوسط الحسابي لعلامات الطالبات في نفس الصف هي 15 فإذا كان عدد الطلاب في الشعبة 15 وعدد الطالبات 10 فإن الوسط الحسابي لعلامات الشعبة كاملة هي ....

أ) 18 ب) 20 ج) 15 د) 17.5

26-إذا علمت أن الوسط الحسابي لعلامات30 طالباً في امتحان مادة الإحصاء هو 70 ، ثم انسحب طالب علامته 80 ، فإن الوسط الحسابي للعلامات بعد الانسحاب هو ....

أ) 70 ب) 69.56 ج) 69.65 د) 80

س3: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- ما قيمة المدى للعلامات التالية 4، 12، 1-، 8، 11:-

أ) 1- ب) 11 ج) 13 د) 7

2- الانحراف المتوسط للقيم 1، 2، 3 هو:-

أ) $\sqrt{2}$ ب) $\sqrt{3}$ ج) $\sqrt{\frac{2}{3}}$ د) 3/2

3- الانحراف المعياري للقيم 5، 4، 3، 2، 1 هو:-

أ) $\sqrt{2}$ ب) $\sqrt{3}$ ج) $\sqrt{5}$ د) $\sqrt{10}$

4- إذا كان الانحراف المعياري لـــــــ (10) قيم هو 9 فإن التباين =

أ) 3 ب) $\sqrt{10}$ ج) 81 د) 100

5- إن مقاييس التشتت تتأثر بالعمليات التالية :

أ) الجمع ب) الطرح ج)الضرب د) لاشيء مما ذكر

* إذا علمت أن درجات 10 طلاب في مادة الإحصاء هي كما يلي:

50 ، 70 ، 85 ، 80 ، 40 ، 65 ، 60 ، 90 ، 85 ،75 أجب عن الأسئلة من 6-10

6- أ ن المدى لدرجات الطلاب هو:

أ) 50 ب) 25 ج)26 د) 40

7- أ ن الانحراف المتوسط لدرجات الطلاب هو:

أ) 130 ب) 13 ج)10 د) 50

8- أ ن الانحراف المعياري لدرجات الطلاب هو:

أ) 900 ب) 10 ج)90 د) $\sqrt{90}$

9- أ ن التباين لدرجات الطلاب هو:

أ) 900 ب) 10 ج)90 د) $\sqrt{90}$

10- أ ن معامل الاختلاف لدرجات الطلاب هو:

أ) $\sqrt{70}$ ب) 70/ $\sqrt{90}$ ج)90 د) $\sqrt{90}$

س4: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14						
الإجابة										

1- حسب معامل الارتباط بين X , Y فكان 0.65- ، عرف المتغيران $L = 2X - 0.2$, $M=-3X+0.3$ فإن معامل الارتباط الجديد بين L , M هو:

أ- (0.35-) ب- (0.85-) ج- (0.65-) د- 0.65

2- إذا كان $\Sigma(x-x')(y-y') = 27-$ ، وكان $\Sigma(x-x')^2 = 36$ ، وكان $\Sigma(y-y')^2 = 81$ ،

فإن معامل ارتباط بيرسون يساوي $r_{xy} =$?

أ- $\frac{(-27)}{36}$ ب- $\frac{(-27)}{81}$ ج- $\frac{(-27)}{9}$ د- $\frac{(-27)}{54}$

3- احدى القيم التالية يمثل معامل ارتباط سلبي عال جداً:

أ- (-0.7) ب- (0) ج- (-0.98) د- (0.99)

4- إذا كانت قيم X كلها 5,5,5,5,5,... وقيم Y كلها -3، -3، -3، -3،... فإن معامل ارتباط سبيرمان=

أ- 1 ب- (-1) ج- 0 د- 0.12

5- إذا أردنا قياس قوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الاسلوب الإحصائي المناسب هو:

أ- تحليل الارتباط ب- تحليل الانحدار ج- التوقع الرياضي د- احتمال الحدث

6- إذا أعطيت بيانات عن متوسط درجات الحرارة وكمية الاستهلاك من الكهرباء خلال 50 يوم من أيام الصيف فإنه يمكن القول أن اتجاه العلاقة بين المتغيرين:

أ- عكسي ب- طردي ج- قوي د- ضعيف

7- نوع الارتباط بين مساحة المربع وأبعاده هو:

أ- تام، بسيط ب- غير تام، بسيط

ج- غير تام، متعدد د- تام، متعدد

8- عند حساب معامل الارتباط بين لون العيون ل (50) طالبة ونسبة الذكاء كان معامل الارتباط (0.9) فإنه يمكن تفسير هذا الارتباط القوي ل:

أ- التأثر بمتغير ثالث فقط ب- التأثر بأكثرمن متغير

ج- العلاقة التبادلية د- عامل الصدفة

9- القيمة التي تمثل أقوى معامل ارتباط عكسي مما يلي هو:

أ- (-2) ب- (-0.7) ج- (-0.90) د- (-1.5)

10- قام طالب بحساب معامل ارتباط بيرسون بين متغيرين فوجد أنه = (1.3) فإن ذلك يدل على:

أ- خطأ في الحساب ب- عدم وجود ارتباط

ج- طردي تام د- عكسي تام

11- إذا كان مجموع مربعات فروق الرتب بين (6 قيم) للمتغيرين X, Y هو (50)، فإن معامل ارتباط سبيرمان للرتب بين X, Y يساوي:

أ- (-1.43) ب- (-0.43) ج- (-0.57) د- (0.43)

12- إذا كان $\Sigma(x-x')\ (y-y') = -5$ ، وكان $\Sigma(x-x')^2 = 9$ ، وكان $\Sigma(y-y')^2 = 4$ ، فإن معامل ارتباط بيرسون يساوي $r_{xy} =$?

أ- (-5)/36 ب- (-5)/4 ج- (-5)/6 د- (-5)/9

13- حسب معامل ارتباط سبيرمان بين X, Y فكان 0.6 فإذا كانت n=10 ، فإن $\Sigma\ d^2$ يساوي:

أ- 44 ب- 66 ج- 0.4 د- 0.66

14- إذا كانت r1=0.6 , r2=0, r3=0.43 , r4= -0.9 فإن معامل الارتباط الذي يعبر عن أقوى علاقة هو:

أ- r1 ب- r2 ج- r3 د- r4

س5: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك.

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- إذا كان Y = 70 + 0.2 X هو خط انحدار Y على X وكانت Y = 72 فإن.... = X

أ- 0.2 ب- 2 ج- 10 د- 20

2- إذا كان Sx = 0.4، Sy=0.2 وكان Y* = 2 + 0.3 X، ما قيمة معامل الارتباط بين Y, X؟

أ- 0.3 ب- 0.4 ج- 0.5 د- 0.6

3- إذا كانت معادلة خط انحدار النفقات على الدخل (X) هي $Y = 0.3\ X + 24$ فإذا كان الدخل 1000 دينار فإن حجم النفقات المتوقع هو:

أ- 324 ب- 276 ج- 1024 د- 976

4- إذا أردنـــا قـــياس قـــوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الاسلوب الإحصائي المناسب هو:

أ- تحليل الارتباط ب- تحليل الانحدار ج- التوقع الرياضي د- احتمال الحدث

5- حـــسب الوسط الحسابي للمتغير X فكان 60، وحسب الوسط الحسابي للمتغير Y فكان 70 فإذا كانت معادلة خط الانحدارYعلىX هي$Y = 0.5X + b$ فإن قيمةb هي:

أ- 30 ب- 40 ج- 60 د- 70

* إذا كـــان معامل الارتباط r_{xy}=0.90، وأن S_x=5.2، S_y=4.3، أوجد معادلة الانحدار y/X علماً بأن قيم x'=12، y'=15 للإجابة عن الأسئلة 6 - 8

6- أن قيمة b هي: أ- 0.744 ب- 6.07 ج- 0.77 د- 6.75

7- أن قيمة a هي: أ- 0.744 ب- 6.07 ج- 0.77 د- 6.75

8- أن معادلة الانحدار $Y = a + bx$ هي:

أ- 0.744+6.07X ب- 0.744+6.07

ج- 6.07+0.744X د- 6.07+0.744

* إذا كان معامل الارتباط r=0.952، وان Sx=5.521 وان Sy=5.062 ، أوجد معادلة انحدار y/x علماً بأن x'=13, y'=16.571 اجب عن الأسئلة 9 – 10

9- إن قيمة a تساوي: أ- 0.952 ب-0.873 ج-52.2 د-5.22

10- إن قيمة b تساوي: أ- 0.952 ب-5.22 ج-52.2 د-0.873

الفصل الثاني

التوزيعات المجتمعية الاحتمالية والتوزيعات العينية

(التوزيع الطبيعي Z، التوزيع التائي T، التوزيع الكائي χ^2، التوزيع الفائي F)

Probability Distributions & Sampling Distributions

الفصل الثاني

التوزيعات المجتمعية الاحتمالية والتوزيعات العينية

Probability Distributions & Sampling Distributions

1-2 مقدمة Introduction

المتغيرات العشوائية Random Variables: أي متغير عشوائي ينتمي إلى عائلة من التوزيعات، هذه العائلة إما معروفة أو غير معروفة.

المتغيرات العشوائية Random Variables: دالة تمثل العلاقة بين فضاء العينة Ω ومجموعة الأعداد الحقيقية R ولها صفات وخصائص محددة.

المتغير العشوائي المنفصل Discrete Random Variables: يأخذ قيم متميزة ويأخذ عدداً محدوداً ومعدوداً من القيم. مثل عدد الطلاب، عدد السيارات.

المتغير العشوائي المتصل Continues Random Variables: يأخذ مجالاً أو حيزاً على خط الأعداد الحقيقية، مثل الطول، الوزن، درجة الحرارة.

*** التوزيعات الطبيعية**

1- عائلة التوزيعات الطبيعية Family of Normal Distributions: أفراد هذه العائلة يتحددوا بمعلمتين هما (الوسط الحسابي، التباين) $N(\mu , \sigma^2)$

2- عائلة توزيعات ذات الحدين Binomial Distributions: أفراد هذه العائلة يتحددوا بمعلمتين هما: عدد المشاهدات، احتمال حدوث النجاح. (عدد المشاهداتn، احتمال حدوث النجاحP) N(n,p)

3- عائلة التوزيعات متعددة الحدود Multinomial Distributions: في هذه التوزيعات هناك معالم $(n, p_1, p_2, \dots , p_k)$، وهو توسيع لذات الحدين.

4- توزيع بويسون Poison Distributions: له معلمة رئيسية واحدة هي m.

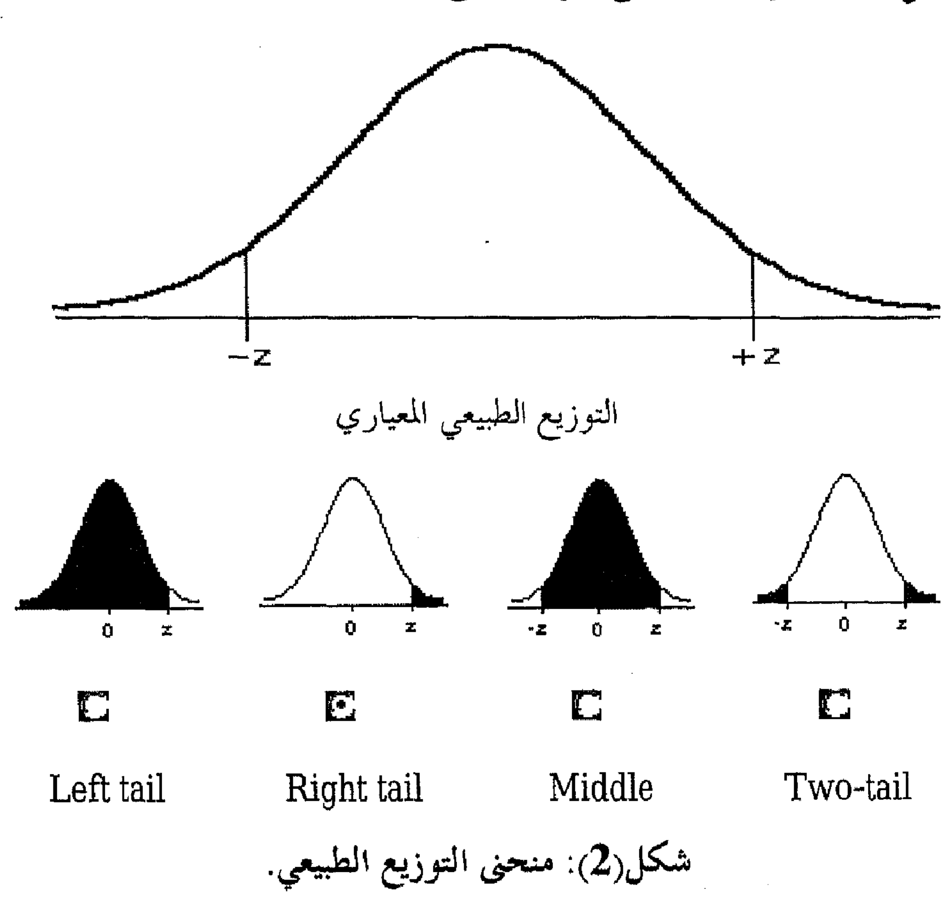
5- التوزيع الاحتمالي لمربعات كاي χ^2 Probability Distribution of χ^2: وشكل توزيعه يعتمد على درجات الحرية df

6- التوزيع الاحتمالي الفائي F Probability Distribution -F: يعتمد على قيمتين لدرجات الحرية $V_1=n_1-1$ و $V_2=n_2-1$

7- التوزيع الاحتمالي التائي (T) Probability Distribution: باستخدام $(x'_i - \mu)/S\,x'_i$ حيث أن $S\,x'_i$ هو تقدير الخطأ المعياري $S_{x'} = S/\sqrt{n}$

2-2 التوزيع الطبيعي Normal Distribution

تمثيل البيانات الإحصائية على صورة منحنى.



شكل(2): منحنى التوزيع الطبيعي.

كثير من الباحثين يؤكدون انه إذا تم اخذ عدد كبير من المشاهدات على إحدى الظواهر مثل الذكاء أو التحصيل الدراسي فإن القيم التي تأخذها هذه المشاهدات تتوزع فيما بينها توزيعاً طبيعياً.

وعلـــى الرغم مما يحيط بهذا الاصطلاح "التوزيع السوي" من غموض وإبهام فلا بد من التعرف على هذا المفهوم بشكل صحيح.

إن المـــنحنى الـــسوي هو التمثيل البياني للتوزيع السوي، وهو حالة خاصة من حالات المـــنحنى الممهـــد المنتظم وهو ممهد ومنتظم تماماً وقائم على عدد كبير من الحالات، فإنه يمكن الحـــصول عليه عن طريق التوزيعات التكرارية المتضمنه بيانات حقيقية وليس عن طريق بيانات افتراضية عن طريق الظواهر ذات الصلة.

خصائص التوزيع الطبيعي:

1. **خاصـــية الاستمرارية Continuity**: عند تمثيل متغير بشكل بياني بمنحنى تكراري فإنه يمكن أن يأخذ أي قيمة ضمن المدى الصفري للتوزيع.

2. **خاصـــية التقارب Asymptotic**: عند تمثيل متغير بشكل بياني بمضلع تكراري فإنه يمكـــن الاصطلاح على فئة صفرية عند كل من طرفي التوزيع ويمكن الحصول على مضلع مغلق بمعنى أن المنحنى يكون تقاربياً.

3. **خاصـــية الـــتماثل Symmetric**: إن نسبة عالية من الأفراد يتجمعون في منطقة المتوســـطات ويـــتوزع الباقون بنسب متفاوتة على جانبي تلك المنطقة، وهذا يعني وجود نقطة انعكاس للمنحنى ويكون ميله عند تلك النقطة يساوي صفراً.

4. **صيغة دالته الاحتمالية:**

$$F(x) = (1/\sqrt{2\pi\sigma^2}) \text{Exp} [-0.5((x-\mu)/\sigma)^2]$$

ويرمز لهذا التوزيع بالرمز $X \sim N(\mu,\sigma^2)$

5. **المساحة تحت المنحنى الطبيعي = 1**، وتقسم المساحة إلى نصفين متساويين عند قيمة $\overline{X}$، ولإيجاد الاحتمالات الخاصة بالتوزيع الطبيعي نستخدم المعادلة:

$$P(a < x \leq b) = {}_a\int f(x)dx$$

العلامات المحولة Transformed Scores

إن المشاهدة الخام غير قابلة للتفسير لأنها ليست كافية لإعطاء صورة عن الوضع الصحيح للمشاهدة، لأنه لا بد من معرفة الوسط الحسابي، وأعلى وأدنى مشاهدة، وأي إحصائي تشتت عن هذه المشاهدات.

أنواع التحويلات:

تصنف الدرجات المحولة تحت نوعين من التحويلات هما:

1- التحويل الخطي Linear

- تحويل المشاهدات من مقياس لآخر بحيث لا يحدث أي تغيير على شكل التوزيع.
- تكون العلاقة بين المشاهدات على المقياسين خطية.
- من أمثلته تحويل العلامة الخام إلى علامة زائية Z score.

2- التحويل غير الخطي Non Linear

- يصاحب تغيير الدرجات تغيير في شكل التوزيع الأصلي.
- تكون العلاقة بين المشاهدات على المقياسين ليست خطية.
- من أمثلته تحويل العلامة الخام إلى علامة معيارية زائية طبيعية Z_{score}.

* الكشف عن نوع التحويل:

- رسم شكل الانتشار. - تمثيل المشاهدات بيانياً.
- إيجاد الفروق بين المشاهدات المتتالية على المقياسين.

* أمثلة على التحويل الخطي:

1- العلامة المعيارية الزائية (Z - Score)

علامـــة معيارية في توزيع وسطه الحسابي = 0 ، وانحرافه المعياري = 1 ، يتم الحصول عليه بالتحويل الخطي دون أي تغيير في شكل التوزيع الأصلي.

$$\text{العلامة المعيارية} = \frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}}$$

$$Z = \frac{X - \bar{X}}{\sigma_X}$$

X : المشاهدة. $\bar{X}$: الوسط الحسابي. σ_X : الانحراف المعياري.

*** خصائص العلامات الزائية:**

- كل مشاهدة يمكن تحويلها إلى مشاهدة زائية.
- الوسط الحسابي للمشاهدات الزائية=0، والانحراف المعياري=1
- تحافظ على شكل التوزيع.

2- العلامة المعيارية التائية (T - Score)

للتخلص من الخصائص غير المرغوبة للعلامات الزائية فإنه يمكن تحويلها إلى أخرى بوسط حسابي وانحراف معياري جديدين، ومن أشهر هذه التحويلات هي العلامة المعيارية التائية بوسط حسابي=50، وانحراف معياري=10 كما تبينه العلاقة أدناه:

$$T = 10 * Z + 50$$

مثال) إذا كانت العلامة المعيارية الزائية = 1، فإن العلامة المعيارية التائية = 60، أما إذا كان التوزيع محدد وسطه الحسابي وانحرافه المعياري مسبقاً، فإن معادلة التحويل تكون:

$$Z' = Z\sigma + \mu$$

*** أمثلة على التحويل غير الخطي:**

1- العلامة الزائية المعدلة (Normalized Z - Score)

علامة معيارية في توزيع وسطه الحسابي = 0 ، وانحرافه المعياري = 1 ، حول شكله الأصلي إلى الشكل الطبيعي.

خطوات التحويل:

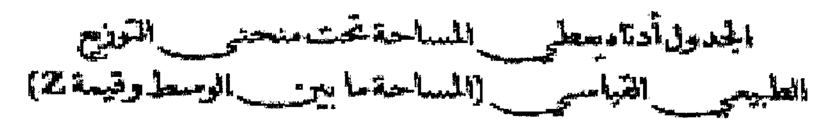
تحويل كل علامة خام في التوزيع الأصلي إلى رتبة مئينية.

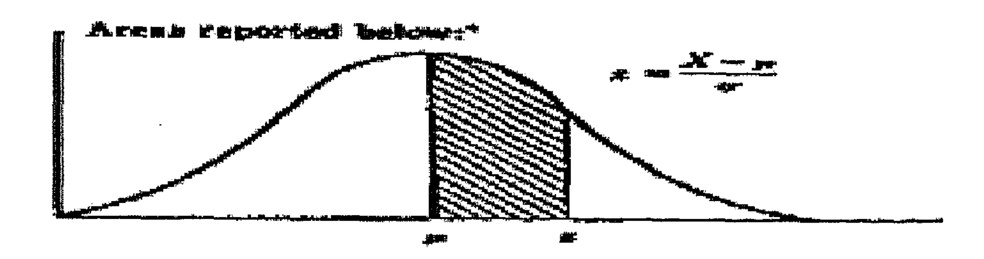
إيجاد العلامة المعيارية الزائية المقابلة لكل رتبة مئينية من جدول التوزيع الطبيعي.

2- علامة ستانين (التساعي) (Stanine Score)

علامة معيارية في توزيع وسطه الحسابي = 5 ، وانحرافه المعياري = 2 ، وأقل علامة في التوزيع هي ستانين1 وأعلى علامة هي ستانين9 وكل ستانين يقابل نسبة معينة من العلامات في التوزيع الأصلي.

التوزيع الطبيعي Normal Distribution


الجدول أدناه يعطي المساحة تحت منحنى التوزيع الطبيعي القياسي (المساحة ما بين الوسط وقيمة Z)



وقـــد تم تحويل قيم التوزيع الطبيعي إلى التوزيع الطبيعي المعياري، وهو توزيع له متوسط حسابيμ =صفر، وانحراف معياريσ =1، ويرمز للمتغير العشوائي المعياري بالحرف Z.

إن الـــتوزيع الطبيعي أكثر التوزيعات الاحتمالية استخداماً لأن توزيع المتغيرات يكـــون طبيعـــياً في أكثر الحالات التطبيقية، كما أنه يمثل تقديراً دقيقاً لعدد كبير من التوزيعات الأخـــرى إذا كـــان عدد المتغيرات كبيراً، وأن لكل توزيع معالم وللتوزيع الطبيعي معلمتان هما الوســـط الحسابي للمجتمع μ والانحراف المعياري للمجتمع σ ، لذلك عرف التوزيع الطبيعي المعـــياري بأنه توزيع له متوسط حسابي=صفر، وانحراف معياري=1، ولكثرة استخدام التوزيع الطبيعي المعياري فقد دون في جداول كما هو موضح في (ملحق رقم 1) ليسهل استخدامه في الحالات التطبيقية، وحدود هذه الجداول هي من (-3 إلى +3)، وإن المساحة الكلية للتوزيع =1 والمساحة فوق Z تساوي صفر = 0.5 ، والمساحة تحت Z تساوي صفر = 0.5

العلامة المعيارية الزائية (Z - Score)

علامـــة معيارية في توزيع وسطه الحسابي = 0 ، وانحرافه المعياري = 1 ، يتم الحصول عليه بالتحويل الخطي دون أي تغيير في شكل التوزيع الأصلي.

$$\text{العلامة المعيارية} = \frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}}$$

$$Z = \frac{X_i - \mu}{\sigma_x}$$

$$Z = \frac{X - \overline{X}}{\sigma_x}$$

حيث:

X_i : المشاهدات.

$\overline{X}$: الوسط الحسابي.

μ : الوسط الحسابي للمجتمع الذي قيمه X_i.

σ_x : الانحراف المعياري للمجتمع الذي قيمه X_i..

*** خصائص العلامات الزائية:**

- كل مشاهدة يمكن تحويلها إلى مشاهدة زائية.
- الوسط الحسابي للمشاهدات الزائية=0، والانحراف المعياري=1
- تحافظ على شكل التوزيع.

خواص المنحنى الطبيعي Normal Distribution Curve

- المساحة الكلية تحت المنحنى=1
- شكل المنحنى على هيئة جرس.
- تتركز المشاهدات حول الوسط الحسابي.
- المنحنى متماثل حول الوسط الحسابي.
- إن المساحة الموجودة على بعد
- ±1 انحراف معياري= 68.27% عن $\overline{X}$
- ±2 انحراف معياري= 95.45% عن $\overline{X}$
- ±3 انحراف معياري= 99.73% عن $\overline{X}$

99-2003- إذا كانت قيمة الوسط الحسابي والوسيط والمنوال متماثلة تماماً في جدول تكراري فإن البيانات يكون لها توزيع يسمى:

أ) **متماثلاً (طبيعياً)** ب) مائلاً إلى اليمين ج) مائلاً إلى اليسار د) مقعراً.

من خصائص التوزيع الطبيعي أن الوسط = الوسيط = المنوال.

مثال 2-1) إذا كانت علامة طالب في مادة الإحصاء هي 75 ومتوسط علامات الإحصاء هو 90 والانحراف المعياري لعلامات الإحصاء هو 5 وعلامته في مادة النظم هي 70 ومتوسط علامات النظم هو 60 والانحراف المعياري لعلامات النظم هو 4 فأي العلامات أفضل.

$$Z_{stat} = \frac{X - \bar{X}}{\sigma_x} = \frac{75 - 90}{5} = \frac{-15}{5} = -3$$

$$Z_{Mis} = \frac{X - \bar{X}}{\sigma_x} = \frac{70 - 60}{4} = \frac{10}{4} = 2.5$$

الذي يحصل على علامة زائية أعلى هو الأفضل.

إذا علامته في مادة النظم MIS أفضل من علامته في مادة الإحصاء.

80-2005- إذا كان الوسط الحسابي لعلامات طلبة احد الصفوف في مادة الإحصاء (60)، بانحراف معياري مقداره (5)، وكانت علامة أحد الطلبة في هذه المادة (75)، فإن قيمة العلامة المعيارية لهذا الطالب هي:

أ- (-3)　　ب- 3/1　　ج- $\sqrt{3}$　　د- 3

$$\text{العلامة المعيارية} = \frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}} = \frac{75-60}{5} = \frac{15}{5} = 3$$

*** المساحة تحت المنحنى الطبيعي**

إن القيمة الاحتمالية لـ Z يتم إيجادها من الجداول الإحصائية للتوزيع الطبيعي، فمثلاً لإيجاد احتمال القيمة 0.3413 نقوم بتحديد موقع القيمة في الجهة اليمين من الجدول ثم نجد القيمة المقابلة لها في الجانب الأيسر وهي 1 وهكذا.

مثال: إذا كان X يمثل توزيعاً طبيعياً بمتوسط $\mu=60$ وانحراف معياري $\sigma=10$ والمطلوب إيجاد الاحتمال التالي:

1. $P(60 \le X \le 65)$　2. $P(65 \le X \le 70)$　3. $P(55 \le X \le 60)$

الحل: 1- $P(60 \leq X \leq 65)$

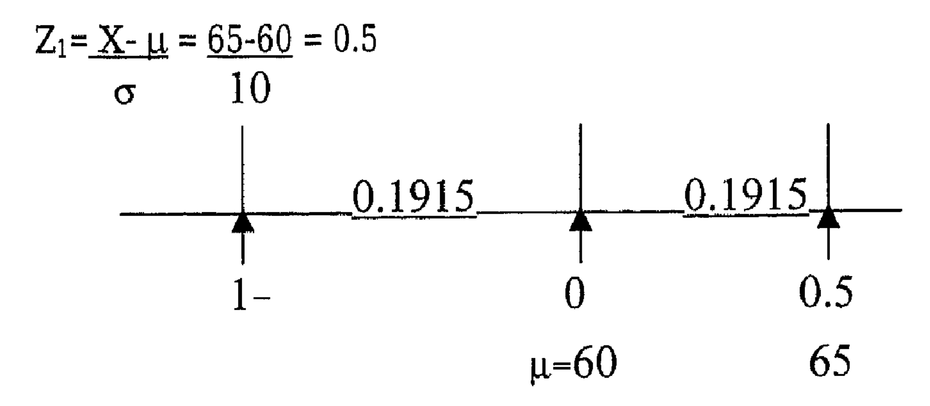
نقوم بتحويل قيم X للتوزيع الطبيعي إلى قيم Z المعيارية باستخدام الصيغة $Z = \frac{X-\mu}{\sigma}$

قيمة Z عند X = 60

$$Z_1 = \frac{X-\mu}{\sigma} = \frac{60-60}{10} = 0$$

قيمة Z عند X = 65

$$Z_1 = \frac{X-\mu}{\sigma} = \frac{65-60}{10} = 0.5$$

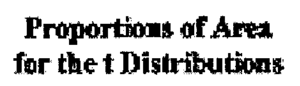


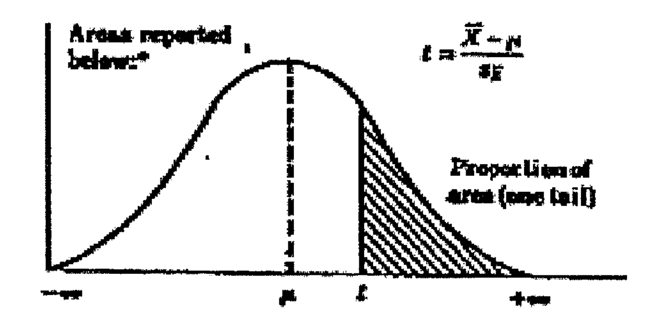
وبالرجوع لجدول التوزيع الطبيعي نجد أن قيمة الاحتمال بين 0-0.5 هي 0.1915

3-2 التوزيع الاحتمالي التائي Probability Distribution (T)

الجدول أدناه يعطي قيمة ∝

المقابلة للمساحة المظللة وقيمتها ∝



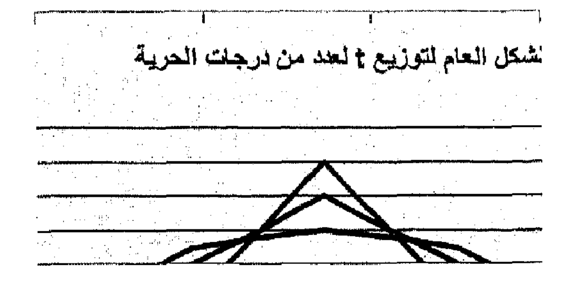


في حالة التوزيع الطبيعي Z إن الانحرافات للمتوسطات العينية (μ - $\bar{X}_i$) وقسمتها على الخطأ المعياري $\sigma_{\bar{X}}$ موزعة توزيعاً طبيعياً مع $\mu=0$ و $\sigma=1$، حيث أن طرح μ من وسط كل

عينة $\bar{x}_i$ لا يغير من شكل التوزيع لمتوسطات العينة وأن قسمتها على $\sigma_{x'}$ يؤدي إلى تقليل التباين إلى 1، ولكن عند حساب الانحراف المعياري لكل متوسط $\bar{x}_i$ باستخدام $(\bar{x}_i - \mu)/S\bar{x}_i$ حيث أن $S\bar{x}_i$ هو تقدير الخطأ المعياري لمتوسط العينة i، يجعل توزيع الانحرافات أكثر وأوسع تسطحاً مما هي عليه في حالة Z، والسبب في هذا أن المقام هو الخطأ المعياري للعينة وليس للمجتمع، وأن هذه الزيادة في الانحراف تكون انعكاساً للتباين الكبير للنسبة $(\bar{x}_i - \mu)/S\bar{x}_i$ ، وأن التوزيع المتوقع لهذه النسبة هو ما يسمى (توزيع t) وصيغة احتسابه كالتالي:

$$t = (\bar{x} - \mu)/S_{x'}$$

حيث أن: $S_{x'} = S/\sqrt{n}$



شكل(3): الشكل العام لتوزيع t لعدد من درجات الحرية.

إن توزيع t له نفس خصائص التوزيع الطبيعيZ حيث انه:متماثل، يمتد من∞- إلى∞+ يختلف عن التوزيع الطبيعي من كون شكله يعتمد على عدد درجات الحرية v حيث أن V=n-1 حيث أن n هي حجم العينة المعتمدة في حساب التباين ويرمز لدرجات الحرية بالرمز df وغالباً ما يقترب توزيع t من شكل التوزيع الطبيعي كلما زادت عدد درجات الحرية وأكثر درجة يتقارب التوزيعان بها عندما تكون df=30 ، وعندما تكون df=∞ يصبح توزيع t توزيعاً طبيعياً، لذلك هناك جدول خاص بتوزيع t يعطي قيمه بناء على قيم df ويوجد هذا الجدول بملاحق الكتاب.

مثال: إذا كان لدينا عينة حجمها n=16 وانحرافها المعياري S=1.5 مسحوبة من مجتمع يخضع لتوزيع t ، وسطه الحسابي $\mu=4$ فما هو احتمال أن تكون قيمة الوسط $\bar{x}=5$

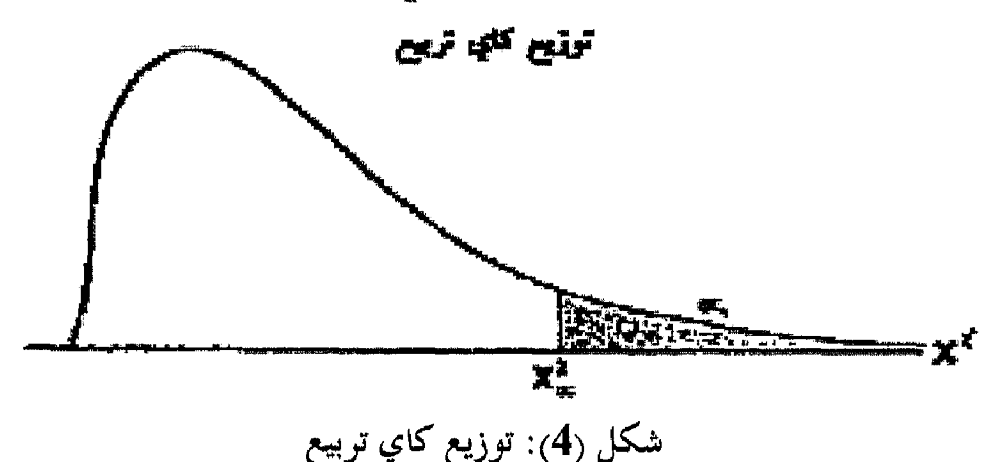
$$S_{x'} = S/\sqrt{n} = 1.5/\sqrt{16} = 1.5/16 = 1.5/4 = 0.375$$

$$t = \frac{\bar{x} - \mu}{S_{\bar{x}}} = \frac{5-4}{0.375} = \frac{1}{0.375} = 2.67$$

t	0	5
	$\mu=4$	2.67

ومـــن جدول t وبدرجة حرية $V=n-1 = 16-1 = 15$ والبحث عن اقرب قيمة لـــ 2.67 المحسوبة نجدها 0.99 وعليه فإن $P(t=2.67) = 0.99$

4-2 التوزيع الاحتمالي لمربعات كاي
Probability Distribution of χ^2



شكل (4): توزيع كاي تربيع

توزيع χ^2 من التوزيعات الاحتمالية النظرية المستمرة، ودالة كثافته النوعية تأخذ القيم من 0 وحـــتى $+\infty$،ولكن شكله لا يشبه توزيع Z ,t لأنه منحنى يقترب للتماثل مع التوزيعات من جانـــبه الأيمن فقط، وشكل توزيعه يعتمد على درجات الحرية df لذلك هو دالة لعدد درجات الحرية ويبدأ بالشكل L ويقترب من التماثل كلما تزداد درجات الحرية.

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma^2}$$

وهي توزيع مع $(n-1)$ من درجات الحرية.

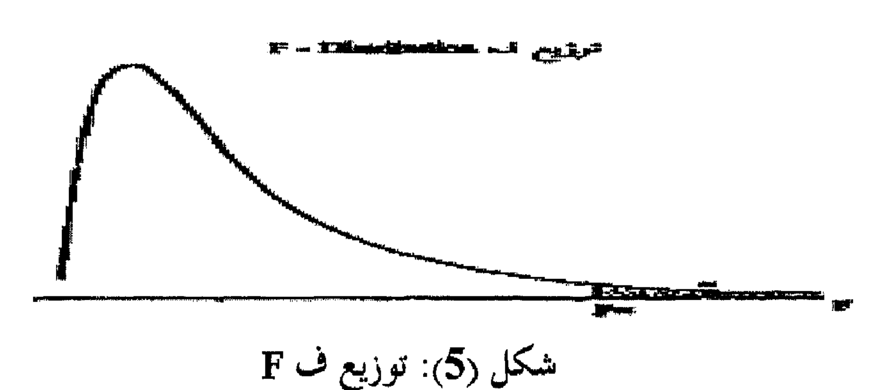
لإيجاد المساحات تحت منحنى χ^2 نستعمل جدول χ^2 حيث يمثل العمود الأيسر درجات الحرية v والخط الأفقي يمثل مستويات المعنوية، والقيم داخل الجدول تمثل المساحات للتوزيع.

مثال: إذا كان لدينا χ^2 المحتسبة هي 0.995 لعينة حجمها n=12 فما هي قيمة χ^2 التي يقع إلى يسارها 0.995 من المساحة.

الحل:

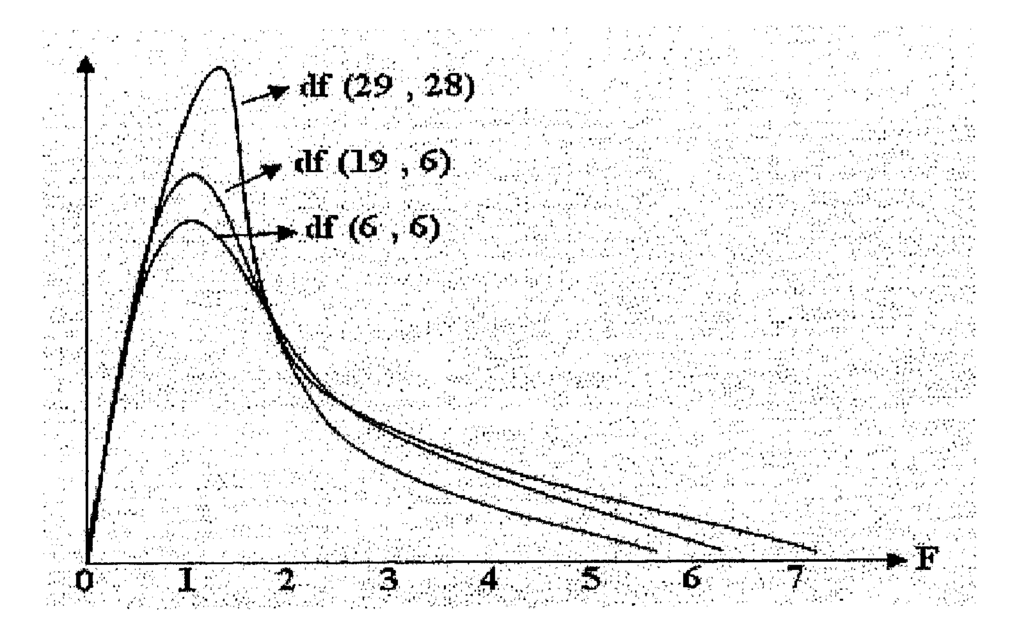
$V = n - 1 = 12\text{-}1 = 11$ من الجدول $\chi^2_{0.995,11} = 26.757$

5-2 التوزيع الاحتمالي الفائي F
F- Probability Distribution



شكل (5): توزيع ف F

توزيع F : هو توزيع ملتو جهة اليمين بمعلمتين تتمثلان بدرجتي حرية (البسط ، المقام) وهمــا k - 1 للبسط ، k – n للمقام حيث n مجموع إحجام العينات، فإذا كان لدينا اختبار لقياس معنوية الفرق بين التقديرين (F) نوجد F_α حيث α مستوى المعنوية المستخدم للفرضية H_0 الــتي ترفض إذا كان $F < F_\alpha$ وإلا نؤكد بوجود الاختلاف بين المتوسطات، والشكل التالي يبين توزيع F



شكل (6): منحنى توزيع F حسب درجات الحرية

كمــا في توزيعات χ^2 , t فإن توزيع F هو من التوزيعات الاحتمالية النظرية المستمرة، ولكــن شــكله لا يــشبه توزيعــي χ^2 , t لأنه يعتمد على قيمتين لدرجات الحرية $V_1=n_1-1$ و $V_2=n_2-1$ وكــل مــنهما تــبدأ من 1 وحتى ∞، ونسبة التباين تعتمد على تباين العينتين S^21/S^22، ونظرياً فإن هذه القيمة تقترب من 1 لأن التباينين هما تقديرات لنفس الكمية ويرمز لها بالرمز Fs.

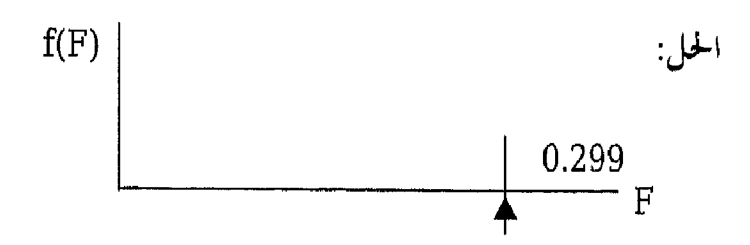
إذا كانت العينتين من مجتمعين طبيعيين منفصلين مختلفين في المتوسطات ولها نفس التباين أي: $\sigma^2_1=\sigma^2_2$ ، $\mu_1\neq\mu_2$ أو أن العينــتين مــسحوبة من نفس التوزيع الطبيعي ولكن تباينهما نفسه، يرمز للتوزيع بالرمز F.

والشكل العام لتوزيع F كما هو شكل توزيع χ^2 يبدأ على شكل حرف L عند درجات الحرية الصغيرة، ثم يصبح مفرطح باتجاه اليمين بازدياد درجات الحرية.

يستخدم الجدول في ملحق الكتاب لإيجاد المساحات تحت المنحنى لتوزيع F

مــثال: جد المساحة تحت المنحنى لتوزيع F عند النقطة 0.95 ودرجات حرية $v_1=10$ و $v_2=8$.

الحل:



ومن جدول F وبدرجتي حرية 8 ,10 حيث $v_1=10$ بالاتجاه الأفقي للجدول و $v_2=8$ بالاتجاه العمودي للجدول و P=0.95 وبذيل واحد نجد القيمة 3.347 فتكون المساحة:

$$F = \frac{1}{F_{(0.95,10,8)}} = \frac{1}{3.347} = 0.299$$

6-2 نظرية المعاينة Sampling Theory

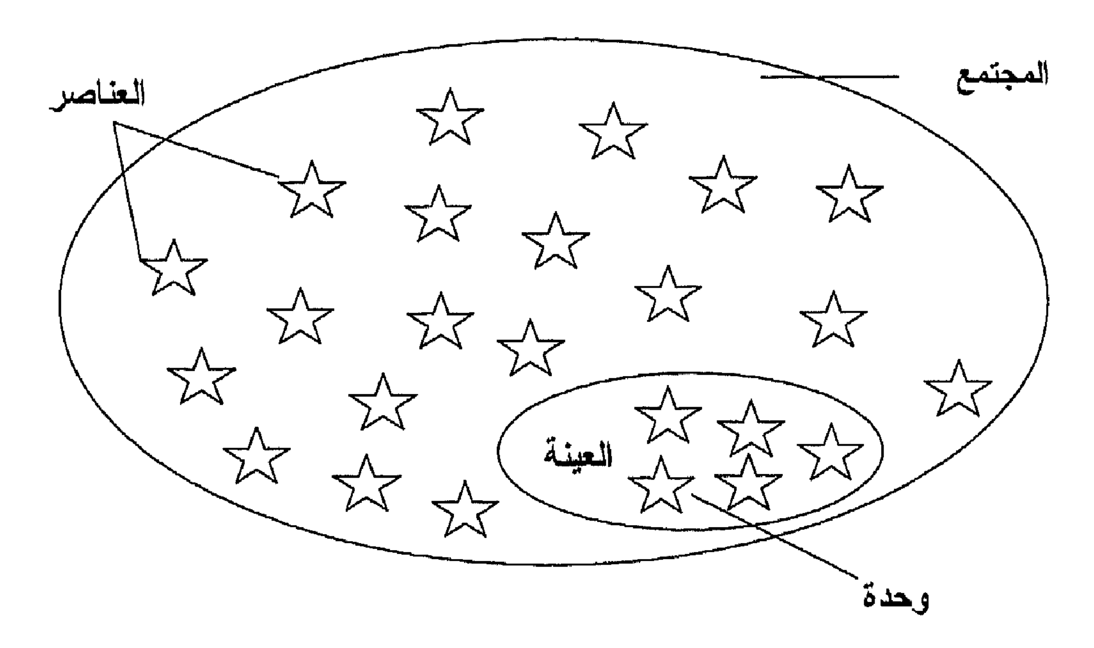
في الاحــصاء الاستدلالي حيث يتوصل الباحث الى خصائص المجتمع Population عن طــريق العيــنة Sample، أو المعاينة، لأن دراسة المجتمع احياناً مستحيلة أو صعبة جداً ومكلفة وتحتاج الى وقت وجهد كبيرين.

المجتمع والعينة والعنصر ووحدة المعاينة

Population, Sample, Element, and Sampling Unit.

إن أسئلة البحث والأهداف التي نحتاجها هي التي تحدد مدى الحاجة إلى استخدام العينة، كمــا أن قيود الوقت والكلفة وطبيعة المجتمع المبحوث قد تحول دون قدرة الوصول إلى المجتمع الكامل لجمع البيانات منه، وهنا لابد من إتباع أحد أساليب المعاينة والتي تزودك بأساليب مختلفة لأنواع العينات.

والشكل التالي يبين العلاقة بين المجتمع والعينة والعنصر ووحدة المعاينة.



الشكل(1): المجتمع والعينة والعنصر ووحدة المعاينة.

Source: Saunders, Mark, Lewis, Philip, & Thornhill, Adrian (2007). ***Research methods for business students*** (4th ed.). Edinburgh Gate, Harlow: Pearson Education Limited. p. 205.

التوزيع Distribution: مجموعة مشاهدات مهما كان عددها.

المجتمع Population: جميع المفردات التي يمكن أن يأخذها المتغير.

مجتمع العينة Sample Population: المجتمع الذي تؤخذ منه العينة.

مجـــتمع الهدف Target Population: المجتمع الذي ستعمم عليه نتائج الدراسة التي أجريت على مجتمع العينة.

العينة Sample: مجموعة جزئية من المجتمع.

المؤشـــر Index: تدل على جميع مقاييس النزعة المركزية والتشتت والعلاقة (الارتباط) سواء محسوبة لعينات أو لمجتمعات.

مؤشرعـــينة (إحصائي) Statistic:ويستخدم للعينات، مثل الوسط الحسابي لعينة X'.

مؤشر مجتمع (مـــعـــلم) Parameter: ويستخدم للمجتمع، مثل الوسط الحسابي μ.

7-2 العينات Samples

المجـــتمع الإحـــصائي Population: أي تجمـــع معـــرف من الأشياء أو الأشخاص أو الحوادث، وهو المجموعة الشاملة التي يجري اختيار العينات منها.

العيـــنة Sample: جزء من المجتمع تتم دراسة الظاهرة عليهم من خلال المعلومات عن هذه العينة، حتى نتمكن من تعميم النتائج على المجتمع.

العيـــنة Sample: أي مجموعة جزئية من المجتمع الإحصائي يتم جمع البيانات من خلالها بصورة مباشرة.

خصائص العينة تختلف باختلاف العينة وحتى يكون التقدير مناسب يجب أن تكون العينة تتمتع بما يلي:

- أن تكون العينة ممثلة للمجتمع.
- أن يكون حجم العينة مناسب.

العينة الممثلة للمجتمع: هي العينة التي يتم اختيارها بطريقة عشوائية.

أسباب استخدام العينات Reasons for using Samples:

1. صعوبة حصر أفراد المجتمعات في وقت واحد.
2. تقليل نفقات الدراسة.
3. صعوبة تأمين العدد الكافي من المختصين الذين تحتاجهم الدراسة.
4. التقليل من الوقت اللازم لإجراء الدراسة والتمكن من تحديده.
5. الحصول على دقة قريبة من استخدامنا للمجتمع.

عدد أفراد العينة Number of Sample's Persons:

1. لا يوجد قانون محدد لتحديد حجم العينة.
2. الدراسات المسحية: 20% من أفراد المجتمع إذا كان صغير نسبياً (500-1000) وتصبح 5% من أفراد المجتمعات الكبيرة جداً.
3. العينة تكون 30 فرداً من أفراد المجتمعات الصغيرة. ولا تقل عن ذلك.
4. الدراسات الارتباطية: 30 فرداً لكل متغير في الارتباط والانحدار المتعددين.

5. البحوث التجريبية: 15 فرداً في كل مجموعة.
6. التحليل العاملي: أن يكون حجم العينة من خمسة إلى عشرة أمثال عدد الفقرات.

ويبين الجدول (1.4.) حجم العينة المطلوب اعتماداً على هامش الخطأ المسموح به، ومنه يتبين أنه كلما قل هامش الخطأ المسموح به زاد حجم العينة.

الجدول (1) اختلاف حجم العينة المطلوب باختلاف هامش الخطأ المسموح به.

حجم المجتمع الكلي	هامش الخطأ المسموح به **(Margin of error)**			
	(%5)	(%3)	(%2)	(%1)
50	44	48	49	50
100	79	91	96	99
150	108	132	141	148
200	132	168	185	196
250	151	203	226	244
300	168	234	267	291
400	196	291	343	384
500	217	340	414	475
750	254	440	571	696
1000	278	516	702	906
2000	322	696	1091	1655
5000	357	879	1622	3288
10000	370	964	1936	4899
100000	383	1056	2345	8762
1000000	384	1066	2395	9513
10000000	384	1067	2400	9595

Source: Saunders, Mark, Lewis, Philip, & Thornhill, Adrian (2007). ***Research methods for business students*** (4th ed.). Edinburgh Gate, Harlow: Pearson Education Limited. p. 212.

كما قدم كل من كريجيسي ومورجان (Krejcie & Morgan, 1970) جدول يسهل اتخاذ قرار جيّد لتحديد حجم العينة المطلوبة اعتماداً على حجم المجتمع الكلي وهامش الخطأ المسموح به (5%).

الجدول (2)

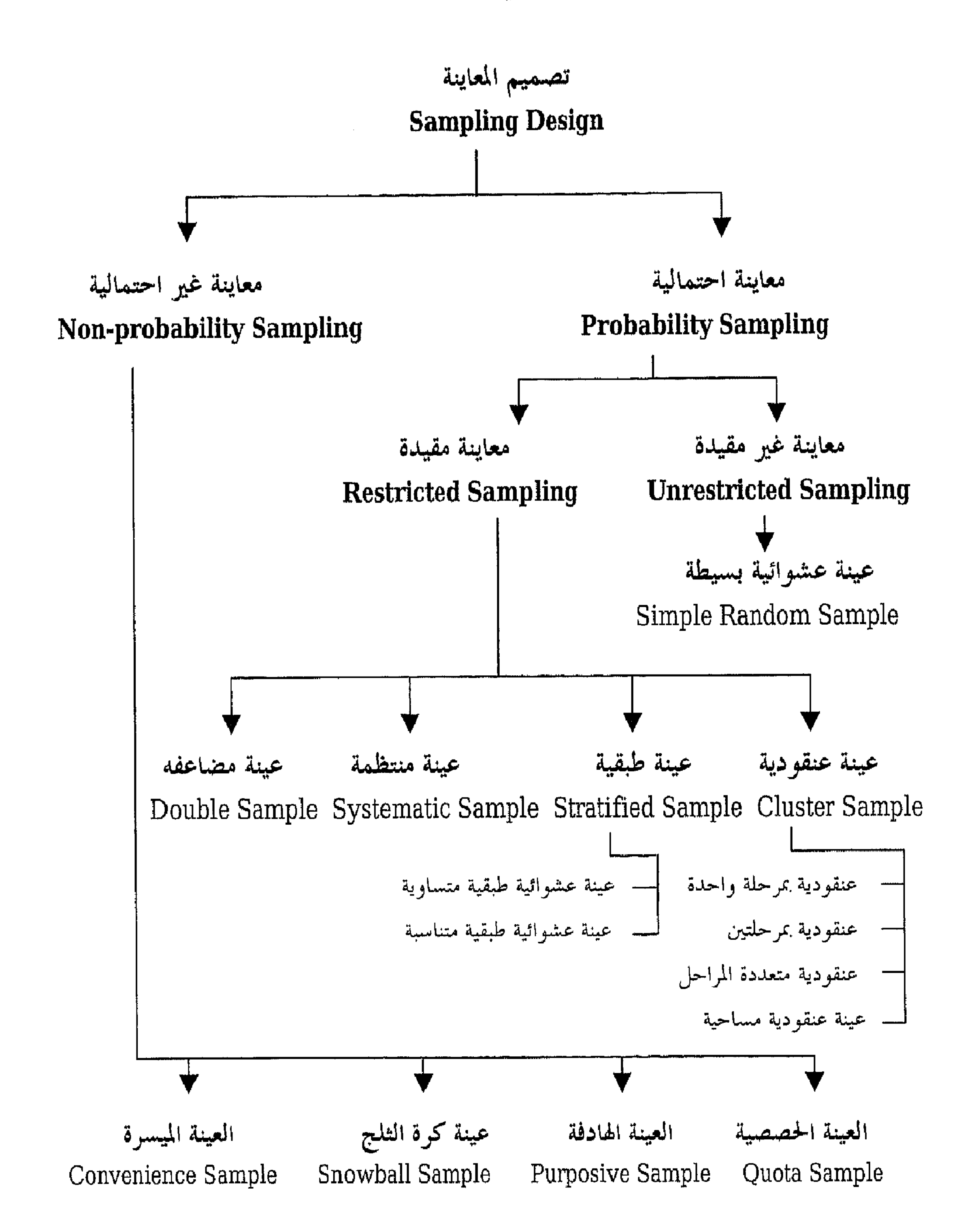
تحديد حجم العينة اعتماداً على حجم المجتمع الكلي (هامش الخطأ المسموح به (5%)).

المجتمع (N)	العينة (n)	المجتمع (N)	العينة (n)	المجتمع (N)	العينة (n)
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

Source: Sekaran, Uma (2003). ***Research methods for business: A skill-building approach*** (4th ed.). New York: John Wiley & Sons Inc., p. 294.

تصنف العينات إلى عينات احتمالية وعينات غير احتمالية:

يبين الشكل (2) الأساليب المختلفة التي يمكن إتباعها في اختيار العينة.



أولاً: العينات غير الاحتمالية Non probabilistic Samples

هـــي عينات يتدخل فيها ميل الباحث وتحيزه بدرجة كبيرة في اختيار أفرادها، ويصعب تعميم نتائجها على جميع أفراد المجتمع. ومنها ما يلي:

1. العينة الميسرة Convenience Sample

تتـــضمن العيـــنة الميسرة اختيار جزافي أو مصادفة للحالات المدروسة والتي من السهولة الحـــصول علـــيها في العينة، إذ يتم اختيار وحدات العينة بناء على سهولة الوصول والاتصال بالأعضاء، وهي سريعة التنفيذ وقليلة الكلفة، ولكن لا يمكن تعميم نتائجها. وغالباً ما تخدم هذه العينة كدراسة قبلية/ أولية أكثر من كونها عينة مهيكلة متكاملة (Saunders et al., 2007, 234).

ومـــثال ذلك إذا أراد الباحث أن يتعرف على رأي الطلبة المبدئي في أداء المواصلات في جامعـــة مـــا فإنه يقوم بسؤال أول (50) طالب أو طالبة يواجههم عند البوابة الرئيسة ليتعرف على آرائهم في أداء المواصلات في تلك الجامعة.

2. العينة الهادفة Purposive Sample

تـــستخدم العيـــنة الهادفة للحصول على معلومات من شريحة مُحدّدة قادرة على توفير المعلـــومات، إمـــا بسبب موقعهم، أو لان بعض المعايير التي وضعها الباحث تتوفر فيهم؛ لأنهم أفـــضل الأشـــخاص القادرين على توفير المعلومات، حيث يتم اختيار وحدات العينة بناء على الخـــبرات في الموضـــوع الذي يدرس. وتستخدم العينة الهادفة عندما تكون المعلومات المطلوبة متوفـــرة لدى فئة معينة من الأفراد، فهي التي تملك المعرفة في الموضوع المبحوث وتستطيع تقديم المعلومة.

وتستخدم العينة الهادفة في الغالب عندما نتعامل مع عينات صغيرة، أو عندما نتعامل مع حالات نريد منها معلومات خاصة (Neuman, 2000).

ومـــثال ذلك: لو أردنا الوصول إلى إجابة السؤال التالي: ماذا يلزم المديرة للوصول إلى مراكز القمة؟ فإن العينة الهادفة المناسبة هنا هي مجموعة النساء التي احتلت مراكز عليا؛ إذ يكون لديهن معرفة متخصصة في ذلك الموضوع نتيجة الخبرة.

3. العينة الحصصية Quota Sample

تستخدم العينة الحصصية في مقابلات المعاينة (Survey) وهي غير عشوائية تماماً، وتقوم على افتراض أن العينة تمثل المجتمع وان التغيّر بالنسبة لمتغيرات العينة الحصصية هي نفسها بالنسبة لمتغيرات المجتمع، لذا فإن العينة الحصصية هي نوع من العينة العشوائية الطبقية ولكنها تختار أفراد الطبقة بطريق غير عشوائي، إذ تعتمد على تقسيم المجتمع إلى مجموعات خاصة، ثم حساب حصة كل مجموعة اعتماداً على علاقتها بالبيانات المتوفرة وحجم المجتمع، ثم الحصول على تلك الحصة بأيسر الطرق (Saunders et al., 2007, 226).

وتستخدم العينة الحصصية عندما يكون هناك صفات محدّدة يجب أن تؤخذ مسبقاً بالاعتبار في العينة مثل: (الجنس، الوظيفة، التوزيع الجغرافي)، إذ لابد والحالة هذه من توزيع العينة بالحصة على المجتمع لتُمثّل التنوّع بداخله.

ومثال ذلك: إذا أردنا توجيه سؤال معين إلى مجموع العمال والعاملات في الشركة المتحدة، وتبين إن العمالة في الشركة تتكون من (30%) من العمال الذكور، بينما نسبة (70%) من العمالة إناث، وتقرر أن يكون إجمالي العينة (10) أشخاص. فإننا سنوجه السؤال إلى أول (3) عمال ذكور، وأول (7) عاملات إناث تتم مواجهتهم في ظروف مريحة وبصورة كيفية دون الاعتماد على الأسلوب العشوائي ليصبح مجموع العينة (3 + 7 = 10) أشخاص.

ومن العوامل التي تشجع على العينة الحصصية توفير الوقت والكلفة والجهد، والحصول على إجابات سريعة من العينة، كما أنها تصبح ضرورة عندما يكون شريحة في المجتمع ذات تمثيل قليل ونرغب في إشراكها في العينة المختارة.

4. عينة كرة الثلج Snowball Sample

تستخدم عينة كرة الثلج عندما نواجه صعوبة في تحديد أعضاء المجتمع المرغوب دراسته، حيث يبدأ الباحث بعينة صغيرة ميسرة، ثم تبدأ العينة بالكبر شيئاً فشيئاً مع سير الدراسة.

وفي هذه الحالة نحتاج إلى الخطوات التالية: (Saunders et al., 2007, 232)

- الاتصال بواحد أو اثنين من حالات المجتمع المرغوب دراسته.
- سؤال هؤلاء لتحديد حالات أخرى يمكن الرجوع إليها لتوفر المعلومات لديهم.

- سؤال الحالات الجديدة لتحديد حالات أخرى جديدة، وهكذا.
- الـــتوقف عندما لا نستطيع الوصول إلى حالات جديدة، أو الوصول إلى حجم عينة مقبول.

ومثال ذلك: إذا أراد باحث أن يدرس تأسيس الإمارة عام 1921 عن طريق المقابلات مـــع الأفراد الذين عايشوا الحدث، ونلاحظ في هذه الحالة أن الأفراد الذين عايشو الحدث ولا زال علـــى قيد الحياة قد يكون عددهم قليل، ولذلك يقوم الباحث بتحديد والاتصال بواحد أو أثـــنين من هؤلاء الأفراد، ثم يقوم بالاستدلال منهم على أفراد آخرين وهكذا حتّى لا يستطيع الوصول إلى أفراد جدد، أو يكون قد استوفى البيانات التي يرغب بجمعها لبحثه.

ثانياً: العينات الاحتمالية Probabilistic Samples:

عيـــنات يتم اختيارها بطرق تعتمد مبادىء الاحتمالات بغرض تمثيل المجتمع، ومنها ما يلي:

1. العينات العشوائية البسيطة Simple Random Samples:

اختـــيار عـــدد معين من أفراد المجتمع بحيث يكون لأي فرد من الأفراد الفرصة نفسها للظهور في هذه العينة، وتستخدم للمجتمع الذي يتكون من عناصر متجانسة.

حجم العينة = نسبة العينة * عدد أعضاء المجتمع

2. العينات العشوائية الطبقية Stratified Random Sample:

يـــتم الحـــصول علـــيها بتقسيم المجتمع الأصلي إلى طبقات أو فئات وفقاً لخاصية معينة كالجـــنس أو مـــستوى التعليم، فإذا كانت عناصر المجتمع غير متجانسة فإننا نقسم المجتمع إلى طبقات، ثم نأخذ عينة عشوائية بسيطة من كل طبقة تتناسب مع حجم الطبقة.

ويمكن تقسيم العينة الطبقية العشوائية إلى:

- **توزيـــع متـــساوي** (Equal Distribution) وهنا تقسم العينة الكلية على الطبقات بالتساوي.

• **توزيـــع متناسب/ نسبي** (Proportional Distribution) حيث يؤخذ عدداً من كل طبقة يتناسب مع حجم الطبقة في المجتمع.

العينة الطبقية = (حجم الطبقة / حجم المجتمع) * حجم العينة

مثال: رأي الآباء والأمهات حول قضية معينة.

89-2003- أردنـا اختـيار عينة طبقية حجمها ن = 100 من مجتمع مكون من 1000 شخص وينقسم إلى طبقتين(400ذكور، 600 إناث)، فإن هذه العينة ستكون مكونة من:

أ)40ذكور، 60إناث ب)60ذكور، 40إناث

ج) 50ذكور، 50إناث د) 30ذكور،70إناث

العينة الطبقية = (حجم الطبقة/ حجم المجتمع) * حجم العينة

حجم الذكور = (1000/100)*400 =40

حجم الإناث=(1000/100)*600 =60

85-2003- في دراسة إحصائية استهدفت طلبة كليات المجتمع، أخذت عينة عشوائية من كل كلية يتناسب عددها مع عدد الطلبة فيها، فإن هذه العينة تسمى:

أ) عنقودية ب) منتظمة ج) معيارية د) **طبقية.**

81-2005- كلية تضم عدة تخصصات مختلفة، يراد اختيار عينة تمثل كل الطلاب في الكلية، فإن أفضل اسلوب لاختيار هذه العينة هو العينة العشوائية:

أ) البسيطة ب) المنتظمة ج) **الطبقية** د) العنقودية.

2-2002- عـدد المهندسـين المـسجلين في نقابـة المهندسين 35000 مهندس و 15000 مهندسة، وأردت اختيار عينة عددها (500) مهندس ومهندسة، فالطريقة الأنسب في اختيار هذه العينة على أساس نقابي هي العينة:

أ- **العشوائية.** ب- المنتظمة. ج- العنقودية. د- الطبقية.

3. العينات العشوائية العنقودية Cluster Random Samples:

عندما نواجه في بعض الدراسات التطبيقية أن وحدات بعض المجتمعات تكون على شكل تجمعات وغالبا ما تكون متشابهة إلى حد كبير بالنسبة للخاصية التي نقوم بدراستها مثل: المدن، الشوارع، الكليات، وغيرها فإن هذه التجمعات عندها تسمى عناقيد (Cluster) إذ يحوي كل عنقود منها على عدد من عناصر المجتمع الأصلية والتي غالبا ما تكون متجانسة، فإننا نلجأ في هذه الحالة إلى العينة العنقودية.

تتميز مجموعات الدراسة المختلفة في المعاينة العنقودية بعدم التجانس بين عناصر كل مجموعة، حيث يوجد اختلافات بين العناصر المشكلة للمجموعة الواحدة، مع وجود تجانس بين المجموعات الجزئية (العناقيد). أي تجانس بين العناقيد ككل، ولكن عدم تجانس داخل العنقود نفسه.

وتقسم العينة العنقودية إلى:

- عينة عنقودية بمرحلة واحدة Single-Stage Cluster Sampling
- عينة عنقودية بمرحلتين Duple-Stage Cluster Sampling
- عينة عنقودية متعددة المراحل Multi-Stage Cluster Sampling
- معاينة مساحية Area Cluster Sampling

4. العينات العشوائية المنتظمة Systematic Random Samples:

وهي نادرة الاستعمال وتتصف بانتظام الفترات بين وحدات الاختيار، أي أن الفرق بين كل اختيار والذي يليه متساوياً في كل الحالات،ويستعمل إذا توفرت قائمة بأسماء أفراد المجتمع فإننا نستطيع اختيار أفراد العينة بحيث يكون الفرد ذو ترتيب معين ضمن أفراد المجتمع ويكون اختيار الفرد الأول من القائمة عشوائياً، مثال: اختيار (100) طالب من أصل (1000) طالب في الجامعة.

نحدد مقدار الفترة=عدد طلاب المجتمع/عدد طلاب العينة = 10=100/1000

تحديد نقطة البدء ويتم اختيارها عشوائياً من 0-9 ولنفترض الرقم 8 هو الاختيار الأول، ونزيد لكل اختيار يليه الرقم 10

5. العينات المعيارية Standard Samples:

عيـــنة تمثل المجتمع الإحصائي تمثيلاً صادقاً وتتفق مقاييسها الإحصائية مع مقاييس المجتمع (الوسط، الوسيط، الانحراف المعياري) ويتم اختيارها بصورة تتابعية.

22-2004- من العينات الاحتمالية العشوائية:

أ) القصدية. ب) الحصصية. **ج) العنقودية.** د) الصدفة.

أخطاء العينات Samples Errors

إن النـــتائج الـــتي نحصل عليها من العينات لا تكون مطابقة تماماً للنتائج في حالة المسح الشامل وذلك لأن نتائج العينات تتعرض لمجموعة من الأخطاء منها ما يلي:

1. **أخطاء عشوائية** (أخطاء الصدفة) **Random Errors**: والسبب في هذا الخطأ هو طريقة اختيار العينة، مثل اختيار (حجم العينة، نوع العينة، تباين عناصر المجتمع).

2. **أخطـــاء التحيـــز Bias Errors**: سببه زيادة أو نقص في البيانات، وقد يحث هذا الخطأ أيضا في المسح الشامل وذلك للأسباب التالية:

أ- الإجابات الخاطئة التي قد يتسبب فيها جامع البيانات.

ب- أخطاء من قبل المستجيب لعدم فهمه السؤال.

ج- أخطاء من قبل المستجيب لأمور شخصية.

د- التحيز في عناصر المجتمع التي تم اختيارها.

ه- عدم الوصول إلى مفردات العينة واستبدالها بأخرى.

و- عدم وجود إطار سليم للعينة.

1-2002- يقصد بالعينة:

أ- المشاهدات التي يتم تطبيقها على جميع أفراد مجتمع الدراسة.

ب- مجموعة جزئية من مجتمع الدراسة.

ج- إحدى وسائل المسح الشامل.

د- طريقة إحصائية في قياس النزعة المركزية.

2002-7- يمكننا الحكم على مدى تمثيل عينة ما للمجتمع المأخوذة منه من خلال:

أ- تجانس أفراد عينة الدراسة.

ب- تمثيل العينة بنسبة تزيد على 10%

ج- بعد أو قرب متوسط العينة عن متوسط مجتمعها مقدراً بوحدات الخطأ المعياري.

د- العينة منتظمة.

2002-8- أفضل نسبة في اختيار عينة الدراسة من مجتمع كبير برأي الاحصائيون:

أ- 2% ب- 4% **ج- 5%** د- 10%

2005-72- العينة الأكثر دقة في تمثيل المجتمع غير المتجانس هي:

أ- العشوائية البسيطة. ب- المنتظمة. **ج-الطبقية.** د-متعددة المراحل.

2005-73- مستوى القياس الذي تستخدم فيه الأرقام بهدف التصنيف فقط هو:

أ- الاسمي. ب- الرتبي. ج- النسبي. د- الفئوي.

2002-92- العيـــنة التي تمنح كل فرد من أفراد المجتمع نفس الفرصة في الاختيار ليكون أحد أفرادها هي العينة:

أ- الطبقية. **ب- العشوائية البسيطة.** ج- المنتظمة. د- متعددة المراحل

2006-93- مستوى القياس الذي تكون وحداته متساوية وليس له صفر مطلق هو:

أ- النسبي. ب- الرتبي. **ج- الفئوي.** د- الاسمي.

2006-100- المقصود بمجتمع الدراسة في الإحصاء:

أ) الأفراد الذين تجرى عليهم الدراسة. ب) العينة التي تقع عليها الدراسة.

ج) الأفراد الذين تعمم عليهم نتائج الدراسة. د) جزء من عينة الدراسة.

8-2 توزيع المعاينة Sampling Distribution

يخضع المجتمع الذي تؤخذ منه العينة لتوزيع معين وهو توزيع المجتمع الاحتمالي لمتغير عشوائي يمثل وحدات ذلك المجتمع، وان التوزيع الاحتمالي للاحصاءة يدعى بتوزيع المعاينة لتلك الاحصاءة والمتمثل بثوابت تعين هذا التوزيع تماماً وتسمى معلمات.

اذا كان المجتمع يخضع لتوزيع طبيعي فإن المعلمين هما الوسط الحسابي μ والانحراف المعياري σ فإذا كانت معلومة فيمكن عندئذ ايجاد جميع الاحتمالات المتوقعة لهذا المجتمع. وكذلك الحال إن كان المجتمع يخضع لتوزيع ذو حدين فإن المعلمة هي احتمال النجاح **(p)**، فإذا علمت **(p)** فإننا نستطيع معرفة توزيعه أي يمكن تحديد مجتمعه، أي أن دالة الكثافة الاحتمالية للمتغير **(x)** والذي يمثل أي فرد من أفراد ذلك المجتمع يكون قد تحدد تماماً، وبعد ذلك يمكن حساب بعض المقاييس عن هذه العينة مثل الوسط الحسابي للعينة الواحدة هو $(\bar{x})$ ويسمى بإحصاءة العينة وهذه القيمة ربما تتغير من عينة إلى أخرى، وأن قيمة الوسط كما ذكرنا فإنها تتغير من عينة إلى أخرى لذا فالمتغير العشوائي هنا هو $(\bar{x})$ أي أن قيمة هذا الإحصاءة $(\bar{x})$ تتغير بتغير العينة، ولهذا يسمى التوزيع الاحتمالي لإحصاءة العينة، ومنها يمكن التوصل إلى تعريف الانحراف القياسي لتوزيع المعاينة والذي يدعى بالخطأ القياس للإحصاءة.

نظرية: إذا كان **(x)** يخضع لتوزيع وسطه μ وتباينه σ^2 وكان $(\bar{x})$ يمثل الوسط الحسابي لعينة حجمها **n** والمسحوبة من هذا المجتمع فإن القيمة المتوقعة لهذا الوسط هو $\mu = \mu(\bar{x})$ أي أن الوسط الحسابي لـــ $(\bar{x})$ هو الوسط الحسابي لجميع الأوساط الحسابية للعينات التي سحبت منها هذه العينات، أما تباين $(\bar{x})$ هو:

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

أي أن تباين هذه العينات يعتمد على تباين المجتمع وعلى حجم العينة، وهو بذلك أقل من تباين المجتمع، وبالتالي كلما كبر حجم العينة قل مقدار الخطأ القياسي للوسط الحسابي $(\bar{x})$ ونقرب وسط تلك العينة من الوسط الحسابي للمجتمع لذا يمكن استخدام تقدير الوسط الحسابي كتقدير لـــ μ، ويجب توفر شرط الإرجاع.

توزيع المعاينة للوسط الحسابي من مجتمع طبيعي:
Sampling Distribution for the Mean of Normal Population:

إذا سـحبت عيـنة عشوائية حجمها (n) من مجتمع كبير له وسط حسابي (5) وتباين معلـوم σ^2 فإنّ توزيع المعاينة للوسط الحسابي $(\bar{x})$ يقترب من التوزيع الطبيعي بوسط حسابي $\mu_{\bar{x}} = \mu$.

والانحراف المقياسي:

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

وأن (z) هـي قيمة المتغير الطبيعي القياسي الذي له وسط حسابي مساوياً صفر وبتباين مقداره واحد.

وهذا ما يسمى بتوزيع المعاينة للوسط الحسابي $(\bar{x})$.

مثال (2):

تخضع علامات الطلاب في مادة الإحصاء لتوزيع الطبيعي بمعدل (70) وانحراف معياري (20)، سحبت منه عينة عشوائية حجمها (36) طالباً، أوجد:

- توزيع المعاينة لهذه العينة.
- احسب احتمال أن يزيد معدل علامات الطلاب عن (78).

الحل:

$$\mu_{\bar{x}} = \mu = 70$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{36}} = \frac{20}{6} = 3.3$$

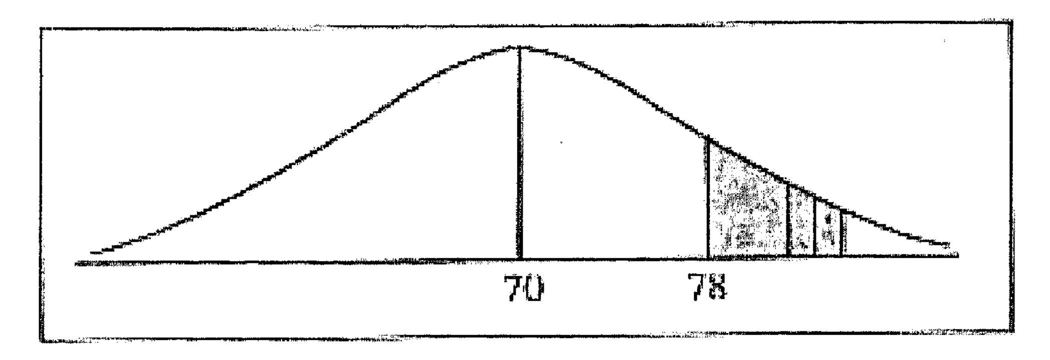
وإنّ توزيع المعاينة هو:

$$\overline{X} \sim N(70,10.89)$$

لإيجاد الاحتمال لدينا:

$$p(\bar{x} \geq 78) = p\left(\frac{\bar{x}-70}{3.3} \geq \frac{78-70}{3.3}\right)$$

$$= p\,(Z \geq 2.42)$$



$$P(Z \geq 2.42) = 0.5 - p(0 < Z < 2.42)$$

وباستخدام جدول التوزيع الطبيعي:

$$P(Z \geq 2.42) = 0.5 - 0.4922 = 0.0074$$

توزيع المعاينة للفرق بين وسطين:

Sampling Distribution for the Difference Between Two Sample Means:

في حالة معلومية تباين المجتمعين الموزعين توزيعاً طبيعياً

نفــرض أن لدينا مجتمعين، الأول بوسط μ_1 وتباينه σ_1^2 والمجتمع الثاني وسطه الحسابي μ_2 وتباينه σ_2^2 والمجتمعين يتبعان التوزيع الطبيعي:

نظــرية: سحبت عينة عشوائية حجمها (n_1) من مجتمع طبيعي معدله μ_1 وتباينه σ_1^2، وعيــنة ثانــية مــن مجتمع طبيعي أيضاً معدله μ_2 وتباينه σ_2^2 والعينة الثانية مستقلة عن العينة الأولى، فإذا كان $(\overline{X}_1)$ يمثل الوسط الحسابي للعينة الأولى، و $(\overline{X}_2)$ يمثل الوسط الحسابي للعينة الثانــية، فإن توزيع المعاينة للفرق بين متوسطيهما $(\overline{X}_1 - \overline{X}_2)$ يتبع التوزيع الطبيعي ذا المعدل $(\mu_1 - \mu_2)$ والتباين:

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

فإنّ توزيع المعاينة يكون:

$$(\overline{X}_1 - \overline{X}_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

يخضع لتوزيع طبيعي معياري بوسط مساوي إلى الصفر وانحراف معياري يساوي واحد، أي أنّ:

$$\mu_{X1-X2} = \mu_1 - \mu_2$$

$$\sigma_{X1-X2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

مثال (3):

سحبت عينيتين عشوائيتين من شركتين مختلفتين لإنتاج العدد الزراعية وكانت الأجور المدفوعة من قِبل الشرطتين تتبع التوزيع الطبيعي، وإن معدل الأجور المدفوعة من قِبَل الشركة (A) إلى (36) عاملاً تساوي ديناراً أردنياً بانحراف معياري (36) ديناراً، أما الأجور المدفوعة من قِبَل الشركة (B) إلى (49) عاملاً تساوي (186) ديناراً بانحراف معياري (40) ديناراً، احسب احتمال أن معدل الأجور المدفوعة من قِبَل الشركة (A) لها متوسط على الأقل (60) ديناراً فأكثر من متوسط الأجور المدفوعة للشركة (B).

الحل:

المعلومات المتوفرة في المثال يمكن توضيحها كما يلي:

الشركة (A)	الشركة (B)	
$\mu_1 = 230$	$\mu_2 = 180$	الوسط الحسابي
$\sigma_1 = 36$	$\sigma_2 = 40$	الانحراف المعياري
$n_1 = 36$	$n_2 = 49$	حجم العينة

$$\mu_{X1-X2} = \mu_1 - \mu_2$$
$$= 230 - 180$$
$$= 50$$

$$\sigma_{X1-X2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= \sqrt{\frac{(36)^2}{36} + \frac{(40)^2}{49}} \doteq 8.3$$

$$P(\overline{X_1} - \overline{X_2} \geq 60) = P\left\{\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq \frac{60-50}{8.3}\right\}$$

$$P\left(Z \geq \frac{10}{8.3}\right) = P(Z \geq 1.204) = 0.5\ P(0 \leq Z \leq 1.204)$$

$$= 0.5 - 0.3925 = 0.1075$$

2- توزيع المعاينة للنسب Sampling Distribution for Proportion:

إذا كانــت قيمة كل عنصر متمثلة بالنجاح أو الفشل فإننا نسمي كل مشاهدة من هذا المجــتمع "تجــربة بيرنوللي" حيث أن احتمال النجاح يساوي (p) واحتمال الفشل يساوي (q) علماً بأن (p + q = 1) وعندما نسحب عينة عشوائية حجمها (n) فإنّه يجب إعادة التجربة (n) مــن المحــاولات، وبذلك فإنّ توزيع المعاينة للمتغير العشوائي (x) المتمثل بعدد النجاحات في العيــنات ذات حجــم (n) يمكــن أن يكــون قريباً من التوزيع الطبيعي بوسط حسابي قدره $\mu = np$ وانحــراف قياسي $\sigma = \sqrt{npq}$ علماً بأن نسبة النجاح مختلفة من عينة إلى أخرى وعلى شرط ألا تكون قريبة من الصفر أو الواحد.

يمكن التعبير عن نسبة النجاح (p) بالمقدار: $\hat{P} = \frac{x}{n}$.

حيث أن (x) = عدد المحاولات (النجاحات).

n = حجم العينة.

$\hat{p}$ = نسبة النجاح في العينة.

وكما ذكرنا أن $\hat{p}$ تختلف من عينة إلى أخرى فإن توزيع المعاينة لـ(x) عدد النجاحات في العينات ذات الحجم (n) يمكن أن يكون قريباً من التوزيع الطبيعي بوسط حسابي قدره:

$$\mu_{\hat{p}} = E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{np}{n} = p$$

وبتباين:

$$\sigma_{\hat{p}}^2 = \text{var}(\hat{p}) = \text{var}\left(\frac{X}{n}\right) = \frac{npq}{n^2} = \frac{pq}{n}$$

على شرط أن تكون قيم (p) قريبة من الصفر أو الواحد.

نظرية:

سحبت عينة عشوائية حجمها (n) من مجتمع إحصائي يخضع لتوزيع بيرنوللي، أي ذات الحدين B(n,p) بوسط $\mu = np$ وتباين $\sigma^2 = npq$، فتوزيع المعاينة إلى $\hat{p}$ نسبة النجاحات هو قريب من التوزيع الطبيعي بوسط $\mu_{\hat{p}} = p$ وانحراف قياسي:

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

لذا فإنّ القيمة المعيارية (Z) هي:

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}}$$

$$\hat{P} = \frac{X}{n}$$

مـــثال (4): إذا كـــان احتمال نجاح الطالب في مادة الاحصاء هو 0.8 وسحبت عينة عشوائية حجمها 49 طالباً، أوجد احتمال **(0.7 < $\hat{p}$ < 0.92)**

الحل: باستخدام النظرية يمكن إيجاد قيم الاحتمال أعلاه كما يلي:

$$P\left(0.7 \le \hat{P} \le 0.92\right)$$

$$P\left\{\frac{0.7 - 0.8}{\sqrt{\frac{(0.8)x(0.2)}{49}}} \le Z \le \frac{0.92 - 0.8}{\sqrt{\frac{(0.8)x(0.2)}{49}}}\right\}$$

$$= P\left\{\frac{-0.1}{0.06} \le Z \le \frac{-0.12}{0.06}\right\}$$

$$= P(-1.66 \le Z \le 2) = P(0 \le Z \le 2) + P(0 \le Z \le 1.66)$$

$$= 0.4772 + 0.4515 = 0.9287$$

توزيع المعاينة للفرق ما بين نسبتين:

Sampling Distribution for Differences Between Two Proportions:

نظرية: سحبت عينتان عشوائيتان حجمها $(n_1, \quad n_2)$ من مجتمعين مستقلين يخضع الأول لتوزيع ذي الحدين $B(n_1, p_1)$، والثانية أيضاً تخضع لتوزيع ذو حدين $B(n_2, p_2)$، وأن:

$$\mu_1 = n_1 p_1 \quad , \quad \sigma_1^2 = n_1 p_1 q_1$$

$$\mu_1 = n_2 p_2 \quad , \quad \sigma_1^2 = n_2 p_2 q_2$$

فتوزيع المعاينة للفرق ما بين $\hat{P}_2 - \hat{P}_1$ يقترب من التوزيع الطبيعي بوسط:

$$\mu_{\hat{p}1-\hat{p}2} = p_1 - p_2$$

وانحراف قياسي:

$$\sigma_{\hat{p}1-\hat{p}2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

لذا فإنّ القيمة المعيارية لهما:

$$Z = \frac{\hat{P}_1 - P_2 - (P_1 - P_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

مثال (5):

إذا كانت نسبة النجاح في مادة الإحصاء في جامعة مؤته تساوي (0.8)، وكانت نسبة النجاح لنفس المادة في جامعة الاسراء تساوي (0.75)، سحبت عينة عشوائية حجمها (70) طالباً من جامعة مؤته وعينة ثانية من جامعة الاسراء حجمها (35).

أوجد احتمال أن تزيد نسبة النجاح في جامعة مؤته عن نسبة النجاح في جامعة الاسراء بمقدار (0.1) على الأكثر.

الحل:

المطلوب هو إيجاد $P\left(\hat{P}_1 - \hat{P}_2 \leq 0.10\right)$.

وباستخدام النظرية أعلاه فإن هذا الاحتمال يساوي:

$$p\ (\hat{p}_1\ \ \hat{p}_2 \leq 0.10)$$

$$= p\left\{\frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \leq \frac{0.10 - (0.8 - 0.75)}{\sqrt{\frac{(0.8)x(0.2)}{70} + \frac{(0.75)x(0.25)}{35}}}\right\}$$

$$= p\ \left(Z \leq \frac{0.05}{0.0873}\right) = p\ (Z \leq 0.573)$$

$$= 0.5 + P\ (0 \leq z \leq 0.573)$$

$$= 0.7157$$

أي أن نسبة النجاح في مؤته تزيد بـــ(71%) عنها في جامعة الاسراء.

توزيع المعاينة للتباين (Sampling Distribution for the Variance):

إذا ســـحبنا عينات عشوائية كل منها ذات حجم n من مجتمع طبيعي تباينه σ^2 ثم أعيد الاختيار لعدة مرات وحسب تباين كل عينة S_i^2 فإننا نحصل على الإحصاءة S^2.

فإن المعاينة S_2 ليس ذا مكانة عملية في الإحصاء لذا نتيجة إلى توزيع المعاينة لـــ:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

هذا يعني أن توزيع المعاينة إلى التباين يخضع لتوزيع مربع كاي وبدرجة حرية $(n\text{-}1)$.

نظـــرية: ســـحبت عينة عشوائية حجمها (n) من توزيع طبيعي معدله (5) وتباينه أي $X \sim N(\mu\sigma^2)$ وكان S^2 هو تباين العينة، لذا فإنّ:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

هي قيمة المتغير العشوائي (χ^2) الذي له توزيع مربع كاي بدرجة حرية $(n\text{-}1)$.

مـــثال: تخـــضع علامـــات الطلاب في مادة القياس والتقويم لتوزيع طبيعي بمعدل 70 وانحراف معياري 10 سحبت منه عينة عشوائية حجمها 25 طالباً.

- جد توزيع المعاينة لهذه العينة.
- احسب احتمال أن يزيد معدل علامات الطلاب عن 78

$\mu_{\bar{x}} = \mu = 70$

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 10/\sqrt{25} = 10/5 = 2$

ان توزيع المعاينة هو:

$\bar{x} \sim N(\mu , \sigma^2/n)$

$\bar{x} \sim N(70 , 100/25)$, $\bar{x} \sim N(70 , 4)$

احتمال أن يزيد معدل علامات الطلاب عن 78

$$P(\bar{x} \geq 78) = P \{ \bar{x} - 70/ 2 \geq 78-70/2 \}$$
$$= p(Z \geq 4)$$

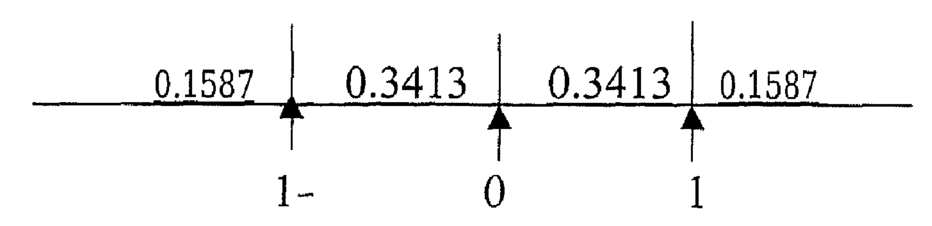
9-2 أسـئلة وتمارين Exercise

س1: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

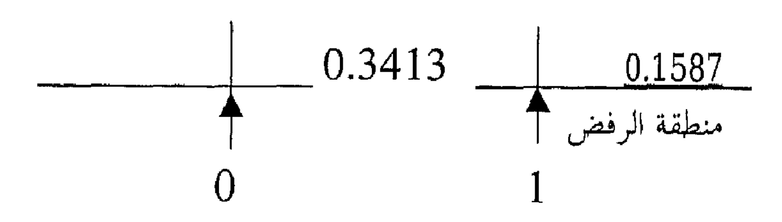
1- في المنحنى الطبيعي المعياري المساحة المحصورة بين Z=-1 و Z=+1 هي:

أ- 0.1587 ب- 0.3413 ج- 0.6826 د- 0.3174



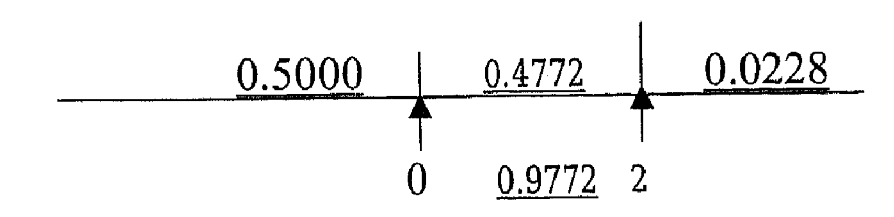
2- إذا كـــان X'=60 ، Sx=10، حيث X علامة الإحصاء، ما احتمال الحصول على علامة أكبر من 70؟

أ- 0.3413 ب- 0.5000 ج- 0.1587 د- 0.8413



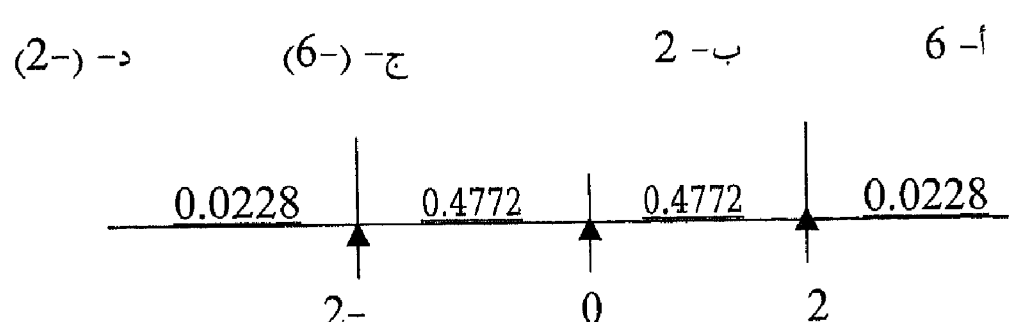
3- تضم جامعة (10000) طالباً وطالبة وتتوزع أوزانهم توزيعاً طبيعياً معيارياً بوسط حسابي X' = 60 وانحراف معياري Sx = 10 ، ما عدد الطلبة الذين تقل أوزانهم عن 80كغم؟

أ- 0.9772 ب- 0.5000 ج- 0.0228 د- 0.4772



4- استعن بالجدول لإيجاد 'X إذا علمت أن X = 16 ، Sx =5:

أ- 6 ب- 2 ج- (-6) د- (-2)



5- إذا تم تحويل علامات من مقياس إلى آخر و لم يصاحب التحويل تغييراً في شكل التوزيع فإن التحويل من النوع..............................

6- العلامـة في توزيـع وسـطه الحـسابي=0، وانحرافه المعياري=1 تسمى العلامة المعيارية ....

7- العلامـة في توزيـع وسطه الحسابي=50، وانحرافه المعياري=10 تسمى العلامة المعيارية ....

* إذا كـان لـتوزيع وسط حسابي=30، وانحراف معياري=5 وكانت لعلامة طالب انحرافً معيارياً واحداً دون الوسط الحسابي فإن:

8- العلامة الخام للطالب = ... ،

9- العلامة التائية للطالب = ...

10- يسعى الباحث إلى تحويل العلامات من توزيع إلى آخر بغرض ....

س2: إذا كانـت درجـات الطلاب في امتحان الرياضيات تتبع التوزيع الطبيعي، وان متوسط الدرجات هو 60 وانحرافها المعياري 10 درجات، ما هو احتمال أن تكون درجة احد الطلاب : الدرجة تقع بين $60 \leq X \leq 70$، الدرجة تقع بين 85 X<50<

س3: إذا كان احتمال رسوب الطالب في مادة الإحصاء هو 0.40 وأخذت عينة حجمها n = 32 طالب من أولئك الذين يدرسون مادة الاحصاء، فما هو احتمال P<=0.30

س4: إذا كانت نسبة الإناث في جامعة اليرموك هي 0.40 ونسبتهم في جامعة مؤته هي 0.60 واخذت عينة من طلبة جامعة اليرموك حجمها 100 وعينة من طلبة جامعة مؤته حجمها 200 فما هي صيغة التوزيع التقريبي للفرق بين نسبتي العينتين مع ايجاد احتمال P[|p2-p1|<=0>1]

س5: اذا كانت الاجور اليومية لعمال احدى الشركات تخضع لتوزيع طبيعي N(5,0.5) وكانت الاجور اليومية لعمال شركة اخرى مماثلة في النشاط تخضع لتوزيع طبيعي N(5,0.25)، واخذت عينة من الشركة الأولى حجمها n1=20 ومن الشركة الثانية حجمها n2=10 أوجد P[x1-x2 <= 0.75]

الفصل الثالث

التقدير وفترات الثقة

تقدير معالم المجتمع من معالم العينة

(باستخدام عينة واحدة)

Estimation of Population Parameters

3-1 مقدمة Introduction
3-2 خواص جودة التقدير
Properties of Goodness of Estimation
3-3 أنواع القيم التقديرية لمعالم المجتمع.
3-4 فترات الثقة Confidence Interval
3-5 العلاقة بين معلمات المجتمع وإحصائيات العينات.
3-6 تمارين Exercise.

الفصل الثالث

التقدير وفترات الثقة

تقدير معالم المجتمع من معالم العينة (باستخدام عينة واحدة)

Estimation of Population Parameters

1-3 مقدمة Introduction

المجتمع Population: مجتمع بيانات أو مشاهدات أو علامات يحدد هويته الباحث، ويمثل جميع عناصر الظاهرة في الدراسة.

العينة Sample: مجموعة جزئية من المجتمع، أو من عناصر الظاهرة.

المؤشر Index: تدل على جميع مقاييس النزعة المركزية والتشتت والعلاقة (الارتباط) سواء محسوبة لعينات أو لمجتمعات.

مؤشر عينة (إحصائي) Statistic: قيمة رقمية تصف خاصية معينة تعود للعينة مثل الوسط الحسابي لعينة $\overline{X}$.

مؤشر مجتمع (معلم) Parameter: قيمة رقمية تصف خاصية معينة تعود للمجتمع الإحصائي وتخص المجتمع، مثل الوسط الحسابي لمجتمع μ .

إن الهدف الأساسي من استخدام العينات هو إيجاد تقديرات تكون ممثلة لمعالم المجتمع Population Parameters عموماً مثل الوسط الحسابي والتباين والنسبة ومعامل الارتباط، وذلك لإمكانية استخدام نظرية الاحتمالات التي تقوم على أساسها هذه العينات العشوائية لدراسة الأخطاء المعيارية للتقديرات ولإيجاد فترة الثقة لقيم معالم المجتمع الحقيقية بالاستناد على نتائج العينة، وعن هذه الطريقة نتمكن من تكوين فكرة عن مدى دقة ومقبولية هذه التقديرات.

العينات Samples

العينة Sample: أي مجموعة جزئية من المجتمع الإحصائي يتم جمع البيانات من خلالها بصورة مباشرة.

2-3 خواص جودة التقدير
Properties of Goodness of Estimation

تتلخص معايير دقة التقدير بما يلي:

1- عـــدم التحيز Unbiasness: وهذا يشير إلا أن تقدير الوسط الحسابي $\overline{X}$ للعينة يعـــود إلى مـــتغير عـــشوائي مـــن حيث الحجم وعملية السحب من المجتمع مستوفياً لشروط العشوائية، أي أن: $E(\overline{X}) = \mu$

حيث أن: μ : متوسط المجتمع الذي سحبت منه العينة. $\overline{X}$: متوسط العينة.

فعند تساوي كلا الوسطين يكون التقدير غير متحيز.

2- التناسق Consistency: وهو قرب قيمة التقدير إلى قيمة المعلمة.

كلما ازداد حجم العينة يقال أن التقدير قد حقق التناسق مع العينة.

3- الكفاءة Efficiency: وتشير إلى التغاير في العينة المستخدمة لأي تقدير، إن التقدير الأقل تباين يكون نسبياً الأكثر كفاءة.

3-3 أنواع القيم التقديرية لمعالم المجتمع.

في الغالـــب نعمد إلى اختيار عينة Sample من المجتمع Population وبعد ذلك نحاول دراســـة هـــذه الظاهرة بالنسبة لأفراد تلك العينة، وبالتالي نعمل على تعميم ما نتوصل إليه من نتائج على كافة أفراد ذلك المجتمع.

تنقسم التقديرات إلى نوعين:

1- التقدير بنقطة Point Estimation

يـــتم التقديـــر هنا بنقطة واحدة أو قيمة واحدة محددة، حيث نقدر معلمة المجتمع بنقطة تحسب من بيانات العينة، مثلاً قيمة الوسط الحسابي $\overline{X}$ باعتبارها تقديراً لمعلمة الوسط الحسابي للمجـــتمع μ، إن تباين العينة S^2 لتقدير تباين المجتمع σ^2 ، حيث أن معلمة التباين من العينة هو عبارة عن تباين متوسط العينة كما يلي:

$$\sigma^2_{x'} = \sigma^2 / n$$

وأن الخطأ المعياري يقدر بالاحصاءة كما يلي:

$$S_{x'} = S / \sqrt{n}$$

مـــثال: إذا كـــان لدينا عينة حجمها $n=5$ وقيم وحداتها هي 3,4,5,1,2 مسحوبة من مجتمع مجهول المعالم.

فإن الوسط الحسابي للعينة $x'=3$ هي نقطة التقدير لقيمة متوسط المجتمع μ غير المعلوم.

وإذا افترضـــنا بـــأن نسبة عدد القيم الزوجية للعينة p هي المقدر لنسبة المجتمع π، فإن $p=2/5=0.4$ هي نقطة التقدير، أي قيمة المقدر لنسبة المجتمع p

ونفس القول عند استخدام S ليدل على التقدير للانحراف المعياري للمجتمع غير المعلوم σ.

2- التقدير بفترة Interval Estimation

إنـــنا لا نتوقع الحصول على المعالم المقدرة بدون أخطاء مهما كانت دقة التقدير، مع أن التقديـــر يـــزداد ثقـــة بزيادة حجم العينة، لذا يجب إعطاء فترة معينة لتوقع وقوع معالم المجتمع داخلها،والمقـــصود هـــنا الفترة التي تشمل قيمة المعلمة المجهولة باحتمال معلوم، وهي ما يسمى بفترة الثقة، ويمكن تقدير μ بفترة لها مقدار ثقة معلوم.

4-3 فترات الثقة Confidence Interval

العلاقة بين معلمات المجتمع وإحصائيات العينات. تحدد كما يلي:

1- تقدير المتوسط الحسابي μ.

قيمة متوسط المجتمع (μ) = المتوسط العام لمتوسطات تلك العينات أي أن:

مـــ = (م$_1$ + م$_2$ + م$_3$ + ... + م$_ك$) / ك

$$\mu = (\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \dots + \bar{X}_k) / k$$

ك: عدد العينات المأخوذة من حجم معين من المجتمع. k

مـــ: متوسط المجتمع. μ — م1: متوسط العينة الأولى. $\bar{X}_1$

م2: متوسط العينة الثانية. $\bar{X}_2$ — م$_ك$: متوسط العينة الأخيرة. $\bar{X}_k$

*** الأســـاليب الإحـــصائية المستخدمة في استخراج معلمات المجتمع اعتمادا على احصائيات العينة:**

1- المتوسط الحسابي Mean

لتقديـــر قـــيمة متوسط المجتمع (μ) من خلال متوسط عينة مأخوذة منه 'X نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة متوسط المجتمع μ مساوية لقيمة متوسط العينة 'X أي أن:

$$\mu = \overline{X}$$

ب- التقدير من خلال مجموعة من القيم.

$$\mu = \overline{X} \pm Z * SE$$

حيث:

μ : متوسط المجتمع.

$\overline{X}$: متوسط العينة.

SE: الخطأ المعياري.

Z : العلامة المعيارية المناظرة لمستوى الثقة المستخدم في عملية التقدير.

مثال: إذا كان متوسط أوزان عينة من الأفراد = 70كغم، وان الخطأ المعياري للوزن = 5كغم، وان مستوى الثقة المستخدم في عملية التقدير 95%، جد متوسط المجتمع.

$$\mu = \overline{X} \pm Z * SE = 70 \pm 1.96 * 5$$

والمعلمـــات الخاصة بالمجتمع غير معروفة في كثير من الأحيان، فإن قيمة الخطأ المعياري يـــندر أن تكون معروفة بشكل مسبق، في هذه الحالة يستعاض عنه بالانحراف المعياري للعينة، وتصبح المعادلة التي يستخرج منها متوسط المجتمع كالتالي:

$$\mu = \overline{X} \pm T * (\sigma/\sqrt{n})$$

حيث:

μ : متوسط المجتمع. $\overline{X}$: متوسط العينة.

T : العلامة المعيارية المعدلة. وتستخرج من جداول التوزيع التائي الخاصة بها.

n : عدد أفراد العينة.

نحتاج لمعرفة T إلى معرفة: مستوى الثقة، حجم العينة، درجات الحرية= ن-1

مستوى الثقة	حجم العينة	درجات الحرية	T
%95	30	29	2.462
%95	25	24	2.492
%95	20	19	2.539

مثال: عينة تتألف من 20 طالب من طلبة الصف الثالث وكان متوسط علاماتهم على اختبار في الرياضيات 80 بانحراف معياري مقداره 5 ، على اعتبار أن مستوى الثقة هو 95 %، جد متوسط أداء طلبة الصف الثالث على هذا الاختبار؟

$$\mu = \overline{X} \pm T * (\sigma/\sqrt{n}) = 80 \pm 2.539 * (5/\sqrt{20}) = 80 \pm 2.539 * (5/4.47)$$
$$= 80 \pm 2.84 = (82.84) , (77.16)$$

اذاً فترة الثقة هي [77.16 – 82.84]

وبما أن قيم متوسطات العينات مختلفة فيما بينها حيث أن بعضها = متوسط المجتمع والآخر مختلف عنه، فإن درجة تشتت قيم هذه المتوسطات حول متوسط المجتمع والتي تسمى الخطأ المعياري وقانونه هو:

الخطأ المعياري $SE = \sigma / \sqrt{n}$

وكلما زاد عدد أفراد العينة كلما قل الخطأ المعياري، وكلما اقتربت قيم متوسطات العينات من قيمة متوسط المجتمع، كلما قل الخطأ المعياري.

1- فترة الثقة للمتوسط Confidence Interval for μ

أ- فترة الثقة لمتوسط مجتمع طبيعي μ تباينه σ^2 معلوم.

Confidence Interval for Population Means (Variance Known)

اختيار عينة عشوائية من مجتمع طبيعي، انحرافه المعياري σ معلوم، فإن متوسط العينة $\bar{X}$ يتبع التوزيع الطبيعي بخطأ معياري $\sigma/\sqrt{n}$، حيث أن n هي حجم العينة.

ولمعرفة ($Z_{(1-\alpha/2)}$) نحتاج الى معرفة مستوى الثقة:

مستوى الثقة	قيمة Z
%95	1.96
%99	2.58

وعند ثقة 95% فإن المساحة التي يقع ضمنها المتوسط الحقيقي هي:

$$1.96*(\sigma/\sqrt{n}) < \mu < 1.96*(\sigma/\sqrt{n})$$

وان المتغير الطبيعي المعياري لتوزيع متوسط العينة Z هو :

$$Z = (\bar{X} - \mu) / (\sigma_{X'})$$

مثال: عند ثقة 95% فإن المنطقة التي يقع ضمنها المتوسط الحقيقي μ تقل بمقدار $1.96*(\sigma/\sqrt{n})$ عن المتوسط الحقيقي وتزداد بمقدار($1.96*(\sigma/\sqrt{n}$

وان المتغير الطبيعي المعياري لتوزيع متوسط العينة Z هو :

$$Z = (\bar{X} - \mu) / (\sigma_{X'}) \Rightarrow Z = (\bar{X} - \mu) / (\sigma/\sqrt{n})$$

حيث أن:

μ : متوسط المجتمع الذي سحبت منه العينة.

$\bar{X}$: متوسط العينة.

$\sigma_{X'}$: الانحراف المعياري للمتوسط ، أي $\sigma/\sqrt{n}$

إذا كانت درجة الثقة 100% فإن فترة الثقة تحسب من المعادلة التالية:

$$\bar{X} - (\sigma/\sqrt{n})(Z_{(1-\alpha/2)}) \leq \mu \leq \bar{X} - (\sigma/\sqrt{n})(Z_{(1-\alpha/2)})$$

مـــثال: مجـــتمع موزع توزيعاً طبيعياً انحرافه المعياري $\sigma=11$ سحبت منه عينة عشوائية حجمها $n=26$ وكان متوسط العينة $x'=48$ احسب تقدير متوسط المجتمع μ بثقة 95%

الحل:

نستخرج قيمة الانحراف المعياري للمتوسط

$$\sigma_{x'} = \sigma/\sqrt{n} = 11/\sqrt{26} = 2.157$$

نستخرج القيمة الجدولية $Z_{(1-\alpha/2)}$

$$Z_{(1-\alpha/2)} = Z_{(1-0.05/2)} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96$$

باستخدام القانون: $\overline{X}-(Z_{(1-\alpha/2)})(\sigma/\sqrt{n}) \leq \mu \leq \overline{X}-(Z_{(1-\alpha/2)})(\sigma/\sqrt{n})$

$$48 - (1.96)(2.157) \leq \mu \leq 48 + (1.96)(2.157)$$

$$43.77 \leq \mu \leq 52.23$$

إن تقدير متوسط المجتمع μ بثقة 95% هي في الفترة [43.77 - 52.23]

ب- فترة الثقة لمتوسط مجتمع طبيعي μ تباينه σ^2 غير معلوم وحجم العينة كبير.

Confidence Interval for Population Means(Variance unknown, Large Sample)

$$\overline{X} - (S/\sqrt{n})(Z_{(1-\alpha/2)}) \leq \mu \leq \overline{X} - (S/\sqrt{n})(Z_{(1-\alpha/2)})$$

مـــثال: سحبت عينة حجمها 64 طالب والمتمثلة بأعمار طلاب الجامعة وكان معدل العمر 20 عاماً والانحراف المعياري للعمر 10، قدر العمر الفعلي μ لمجتمع الطلاب بحدود ثقة 95%.

باستخدام القانون: $\overline{X}-(Z_{(1-\alpha/2)})(S/\sqrt{n}) \leq \mu \leq \overline{X}-(Z_{(1-\alpha/2)})(S/\sqrt{n})$

نستخرج القيمة الجدولية $Z_{(1-\alpha/2)}$

$$Z_{(1-\alpha/2)} = Z_{(1-0.95/2)} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96$$

$$20 - 1.96 * (10/\sqrt{64}) \leq \mu \leq 20 + 1.96 * (10/\sqrt{64})$$

$$20 - 1.96 * 1.25 \leq \mu \leq 20 + 1.96 * 1.25$$

$$20 - 2.45 \leq \mu \leq 20 + 2.45 \Rightarrow 17.55 \leq \mu \leq 22.45$$

إن تقدير متوسط المجتمع μ بثقة 95% هي في الفترة [17.55 - 22.45]

ج- فترة الثقة لمتوسط مجتمع طبيعي μ تباينه σ^2 غير معلوم وحجم العينة صغير.

Confidence Interval for Population Means(Variance unknown, Small Sample)

الخطأ المعياري التقديري لمتوسط العينة $S_{X'} = S/\sqrt{n}$

$$t = (\bar{X} - \mu) / (S_{X'}) = (\bar{X} - \mu) / (S/\sqrt{n})$$

$$\bar{X} - t_{(1-\alpha/2),v} * (S/\sqrt{n}) \leq \mu \leq \bar{X} + t_{(1-\alpha/2),v} * (S/\sqrt{n})$$

د- فترة الثقة لمتوسط مجتمع طبيعي μ تباينه σ^2 غير معلوم وتوزيعه غير معلوم.

الخطأ المعياري التقديري لمتوسط العينة $S_{X'} = S/\sqrt{n}$

$$Z = (X' - \mu) / (\sigma_{X'})$$

$$\bar{X} - Z_{(1-\alpha/2)} * (S/\sqrt{n}) \leq \mu \leq \bar{X} + Z_{(1-\alpha/2)} * (S/\sqrt{n})$$

في حالة العينات التي تؤخذ من مجتمع والتي لها جميعها نفس عدد الأفراد فإن العلاقة بين معلمات المجتمع وإحصائيات العينات يمكن أن تحدد كما يلي:

2- قيمة تباين المجتمع (σ^2) = المتوسط العام لتباينات تلك العينات أي أن:

$$\sigma^2 = (\sigma^2_1 + \sigma^2_2 + \sigma^2_3 + \ldots . \sigma^2_k) / k$$

ويكون صحيحا فقط إذا حسب تباين المجتمع من القانون:

$$\sigma^2 = \sum (X - \mu)^2 / N$$

N: عدد أفراد المجتمع.

وحسب تباين العينة من القانون:

$$\sigma^2 = \sum (X - X')^2 / n-1$$

n: عدد أفراد العينة.

التباين Variance:

لتقديـــر قيمة تباين المجتمع σ^2 من خلال تباين عينة مأخوذة منه S^2 نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة تباين المجتمع σ^2 مساوية لقيمة تباين العينة S^2 أي أن:

$$S^{22} = \sigma$$

ب- التقدير من خلال مجموعة من القيم. وتعتمد على:

مستوى الثقة = 1 – مستوى الدلالة = $1 - \alpha$

حجم مستوى الثقة المطلوب، قيمة الإحصائي (χ^2 / v)

وتقدر قيمة تباين المجتمع من خلال العلاقة التالية:

$$\frac{S^2}{\chi^2(\frac{1-\alpha}{2}) / v} \le \sigma^2 \le \frac{S^2}{\chi^2(\frac{\alpha}{2}) / v}.$$

مثال: أخذت عينة عدد أفرادها 25 من بين الإناث في المرحلة الجامعية ووجد أن تباين أطـــوالهن هو 25 ، فمل هو تباين المجتمع في مثل هذه الحالة إذا كان مستوى الثقة المطلوب هو %99

الحل:

$$S^2 = 25$$
$$V = 25 - 1 = 24$$
$$\alpha = 1 - 0.99 = 0.01$$
$$\chi^2(\frac{\alpha}{2}) / v = 0.412$$
$$\chi^2(\frac{1-\alpha}{2}) / v = 1.9$$
$$\frac{25}{1.9} \le \sigma^2 \le \frac{25}{0.412}.$$
$$13.16 \le \sigma^2 \le 60.7$$

3- تقدير فترات الثقة للنسب في مجتمع Interval Estimation of Proportion

في حالة المجتمع: يرمز لنسبة وقوع الظاهرة في المجتمع ب P وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في المجتمع ب Q وهي نسبة الفشل، وهي (1-P).

أما في حالة العينة: يرمز لنسبة وقوع الظاهرة في العينة ب p' وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في العينة ب P وهي نسبة الفشل.

لتقديـــر قـــيمة نسبة المجتمع P من خلال نسبة عينة مأخوذة p منه نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتـــبر قـــيمة وجـــود النسبة لظاهرة ما في المجتمع P مساوية لقيمة وجودها في العينة المأخوذة من المجتمع p أي أن:

$$p = P$$

ب- التقدير من خلال مجموعة من القيم.

تقدير قيمة النسبة في المجتمع ضمن مجموعة من القيم.

$$P = (k/n) \pm Z * \frac{\sqrt{(k/n)*[(1-(k/n)]}}{\sqrt{n}}$$

$$\hat{p} - (\sqrt{pq/n})(Z_{(1-\alpha/2)}) < P < \hat{p} + (\sqrt{pq/n})(Z_{(1-\alpha/2)})$$

P: النسبة في المجتمع. k/n: النسبة في العينة.

n: عدد أفراد العينة. k: عدد الحالات التي تنطبق عليها الظاهرة المدروسة.

Z: تعتمد على قيمة مستوى الثقة المطلوب.

مـــثال: أخذت عينة تتكون من 100 طالب جامعي من الذكور وسئلوا عن رأيهم في حـــرية الاخـــتلاط في الحرم الجامعي، فأجاب 70 منهم بالإيجاب، فما هي النسبة العامة لتأييد الذكور لحرية الاختلاط في الحرم الجامعي، إذا كان مستوى الثقة المطلوب هو 99%.

الحل:

n: 100   k: 70   Z: 2.576

$$0.70 = 100/70 = k/n$$

$$P = (k/n) \pm Z * \sqrt{\frac{(k/n)*[(1-(k/n))]}{n}} = 0.70 \pm 2.576 * \sqrt{\frac{(0.70)*[0.3]}{100}}$$

$$= 0.70 \pm 2.576 * 0.046 = 0.70 \pm 0.12 = 0.82 \;,\; 0.58$$

4- معامل الارتباط Correlation Coefficient

ما المقصود بعبارة "إن معامل الارتباط بين متغيرين كان ذا دلالة إحصائية"؟

إن هـــذه العبارة استدلال إحصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$ وهي تعـــني أن معامـــل الارتـــباط في المجتمع بين المتغيرين لا يختلف عن صفر، ورفض هذه الفرضية الـــصفرية يعني أن قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً، أي أنها لم تنشأ عن خطأ المعاينة.

يستخدم توزيع المعاينة لمعامل الارتباط بيرسون r في اختبار الفرضية الصفرية $H_0: \rho = 0$ لتقديـــر قـــيمة نسبة المجتمع P من خلال نسبة عينة مأخوذة p منه نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتـــبر قـــيمة معامل الارتباط لظاهرة ما في المجتمع r مساوية لقيمة وجودها في العينة المأخوذة من المجتمع r_1 أي أن:

$$r_1 \; r =$$

ب- التقدير من خلال مجموعة من القيم.

تقدير قيمة معامل الارتباط في المجتمع ضمن مجموعة من القيم.

$$Z(r) = Z(r_1) \pm Z * (1/\sqrt{n-3})$$

حيث:

r : معامل ارتباط المجتمع .

r_1 : معامل ارتباط العينة.

n: عدد أفراد العينة.

Z: القيمة المعيارية المناظرة لمستوى الثقة المطلوب والمستخرجة من الجدول.

Z(r): القيمة الانحرافية لمعامل الارتباط (r) والمستخرجة من الجدول.

مـــثال: إذا كان معامل الارتباط بين الذكاء والتحصيل عند طلبة الجامعة هو 0.62 في حالـــة عينة أفرادها 67 طالباً وطالبة، فإن قيمة معامل الارتباط بين الذكاء والتحصيل في هذه الحالة الدراسية بوجه عام إذا كان مستوى الثقة المطلوب هو 0.95

الحل:

r: معامل ارتباط المجتمع .

0.62:r

67 :n

1.96 :Z

$Z(r_1)$: = Z(0.62) = 0.73 من الجدول

$$Z(r) = Z(r_1) \pm Z * (1/\sqrt{n-3}) = 0.73 \pm 1.96*(1/\sqrt{67-3}) = 0.73 \pm 1.96*0.125$$
$$= 0.73 \pm 0.245 = 0.98 \;,\; 0.48$$

من الجدول نجد أن r التي تناظر Z(r)= 0.98 هي 0.7531 أو 0.75

من الجدول نجد أن r التي تناظر Z(r)= 0.48 هي 0.4462 أو 0.45 وبذلك تكون قيم الارتباط المطلوب هي: r=0.75 كحد أعلى ، r=0.45 كحد أدنى.

جدول(40): قيم Z الجدولية الموزعة توزيعاً طبيعياً N(0.1) عند مستويات معنوية مختلفة.

P	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.50	0.000	0.025	0.050	0.075	0.100	0.126	0.151	0.176	0.202	0.228
0.60	0.253	0.279	0.305	0.332	0.358	0.385	0.412	0.44	0.468	0.496
0.70	0.524	0.553	0.583	0.613	0.643	0.674	0.706	0.739	0.772	0.806
0.80	0.842	0.878	0.915	0.954	0.994	1.036	1.080	1.126	1.175	1.227
0.90	1.282	1.341	1.405	1.476	1.555					
P	**0.000**	**0.001**	**0.002**	**0.003**	**0.004**	**0.005**	**0.006**	**0.007**	**0.008**	**0.009**
0.95	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.96	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.97	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.98	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290
0.99	2.326	2.366	2.409	2.457	2.512	2.576	2.652	2.748	2.878	3.090

5-3 تمارين Exercise.

س(1) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- أي من التالية ليست من معايير دقة التقدير:

أ- عدم التحيز. ب- التناسق. ج- الموضوعية. د- الكفاءة.

2- إن التناسق كمعيار من معايير دقة التقدير يعني:

أ- التغاير في العينة المستخدمة لأي تقدير. ب- قرب قيمة التقدير إلى قيمة المعلمة.

ج- $E(\bar{X}) = \mu$ د- تساوي كلا الوسطين للمجتمع والعينة.

3- مجتمع موزع توزيعاً طبيعياً انحرافه المعياري σ=11 سحبت منه عينة عشوائية حجمها n=26 وكان متوسط العينة x'=40 احسب تقدير متوسط المجتمع μ بثقة 95%

أ- [35.77 , 44.23] ب- [35.77 - 37.23]

ج- [29 - 51] د- [26 - 37]

4- عينة تتألف من 25طالب من طلبة الصف السادس وكان متوسط علاماتهم على اختبار في العلـــوم 80 بانحـــراف معياري مقداره 5 ، على اعتبار أن مستوى الثقة هو 95%، جد متوسط أداء طلبة الصف السادس على هذا الاختبار؟

مستوى الثقة	حجم العينة	درجات الحرية	T
%95	30	29	2.462
%95	25	24	2.492
%95	20	19	2.539

أ- [74 -86]
ب- [75 - 85]
ج- [74.508 - 80.492]
د- [77.508 - 82.492]

5- أخـــذت عينة تتكون من 100 طالب جامعي من الذكور وسئلوا عن رأيهم في التدخين، فأجاب 60 منهم بالإيجاب، فما هي النسبة العامة لتأييد الذكور للتدخين، إذا كان مستوى الثقة المطلوب هو 99%.

أ-[0.400-0.700]
ب-[0.474-0.726]
ج-[0.450-0.702]
د- [0.440 - 0.720]

6- ما المقصود بعبارة "إن معامل الارتباط بين متغيرين كان ذا دلالة إحصائية"؟

أ- استدلال إحصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$

ب- تعني أن معامل الارتباط في المجتمع بين المتغيرين لا يختلف عن صفر.

ج- رفـــض هـــذه الفرضية الصفرية يعني أن قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً.

د- جميع ما ذكر صحيح.

7- إذا كـــان معامل الارتباط بين الطول والوزن عند طلبة المدرسة الثانوية هو 0.62 في حالة عيـــنة أفرادها 67 طالباً وطالبة، فما قيمة معامل الارتباط بين الطول والوزن في هذه الحالة الدراسية بوجه عام إذا كان مستوى الثقة المطلوب هو 0.95

أ- [0.45 - 0.75]　　　ب- [0.43 - 0.73]

ج- [0.47 - 0.77]　　　د- [0.54 - 0.57]

س2: أخـــذت عيـــنة عددها 50 عضو هيئة تدريس في الجامعات الأردنية وأعطوا فحصاً في مهـــارات الحاســـوب فكان متوسطهم 70 بانحراف معياري مقداره 2، أوجد المتوسط والتـــباين لمستوى المعرفة بمهارات الحاسوب عند أعضاء هيئة التدريس إذا كان مستوى الثقة المطلوب هو 95%.

س3: إذا كـــان متوســـط تحـــصيل عينة من طلبة الثانوية العامة عددها 100 طالب هو 75 بانحـــراف معياري مقداره 5 ، أوجد من ذلك متوسط علامات الثانوية العامة وتباينه في حالة الثانوية العامة إذا كان مستوى الثقة المطلوب هو 95%.

س4: في امـــتحان كفاءة ضم 500 طالباً، إذا كان عدد الناجحين فيها 300 طالباً، جد من ذلـــك النـــسبة العامة للنجاح في مثل هذا النوع من الامتحانات إذا كان مستوى الثقة المطلوب 95%.

س5: إذا كـــان معامل الارتباط بين المستوى الاقتصادي والذكاء في عينة قوامها 50 فرداً هو 0.8 ، فجد من ذلك قيمة الارتباط بين هذين المتغيرين بوجه عام إذا كان مستوى الثقة المطلوب 95%.

الفصل الرابع

اختبار الفرضيات

Hypothesis Testing

4-1 اختبار الـفـرضـيـــات Hypothesis Testing

4-2 اختبار الفرضيات الإحصائية

Testing Statistical Hypothesis

4-3 مفاهيم أساسية في فحص الفرضيات

Basic Concepts in Hypothesis Testing

* الخطأ من النوع الأول Type One Error.
* الخطأ من النوع الثاني Type Two Error.
* مستوى الدلالة Level of Significance.
* قـوة الاختبار Power of the Test.

4-4 خطوات اختبار الفرضيات Hypothesis Testing Steps

4-5 اسـتخدام برمجـيـة SPSS في حل المسائل.

4-6 تـمـاريـــــن Exercise.

الفصل الرابع

اختبار الفرضيات
Hypothesis Testing

1-4 اختبار الفرضيات Hypothesis Testing

احد أساليب عمل الاستنتاجات الإحصائية والاستدلال الإحصائي والهدف منه الوصول إلى قرار بشأن معلمة المجتمع Parameter من خلال تقديرها من العينة Statistic المأخوذة من ذلك المجتمع.

الفرضية العلمية والفرضية الإحصائية Scientific and Statistical Hypothesis

الفرضية العلميـــة Scientific Hypothesis

حل مقترح لمشكلة يصاغ بشكل استنتاجي كتخمين ذكي يستند على معلومات علمية سابقة. وتتقرر صحة الفرضية العلمية أو خطؤها في ضوء الخبرة والتجربة.

الفرضية الإحصائية Statistical Hypothesis

ادعـــاء أو تصريح حول معلم غير معلوم تخضع للاختبار الإحصائي الذي يحدد قبولها أو رفضها.

الفرضية الإحصائية التي تخضع للاختبار الإحصائي تسمى بالفرضية الصفرية (H_0).

الفرضية الصفرية(H_0) Null Hypothesis: وتصاغ عادة بالنفي.

H_0: هي الهدف المطلوب اختباره وقبولها يعني أن العينة متوافقة مع الفرضية.

واختبار الفرضيات هو اجراء احصائي يستخدمه الباحث لأختبار الفرضية الصفرية ليتبين فيما اذا كانت صائبه ام خاطئه.

أمثلة:

تعني أن متوسط المجتمع لا يختلف عن 70 $H_0: \mu = 70$

تعني أن تباين المجتمع لا يختلف عن 70 $H_0: \sigma^2 = 70$

تعني أن نسبة المجتمع لا تختلف عن 0.03 $H_0: P = 0.03$

معامل الارتباط بين المتغيرات في المجتمع لا يختلف عن0.7 $H_0: \rho = 0.7$

إن الفرق بين متوسطي المجتمعين لا يختلف عن صفر $H_0: \mu 1 - \mu 2 = 0$

وسط المجتمع الأول لا يختلف عن وسط المجتمع الثاني $H_0: \mu 1 = \mu 2$

إن الفرق بين تباين المجتمعين لا يختلف عن صفر $H_0: \sigma^2_1 - \sigma^2_2 = 0$

تباين المجتمع الأول لا يختلف عن تباين المجتمع الثاني $H_0: \sigma^2_1 = \sigma^2_2$

الفرضية البديلة (H_1) Alternate Hypothesis

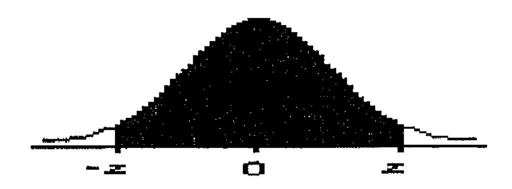
تقبل إذا رفضت الفرضية الصفرية وترفض إذا فشلنا قي رفض الفرضية الصفرية.

تعني أن وسط المجتمع يختلف عن 70 $H_1: \mu \neq 70$

تعني أن تباين المجتمع يختلف عن 70 $H_1: \sigma^2 \neq 70$

تعني أن نسبة المجتمع تختلف عن 0.03 $H_1: P \neq 0.03$

الفرضية البديلة غير المتجهة Non directional Alternate Hypothesis



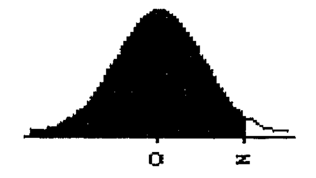
تكـــون محايدة وتنص على انه إذا لم تكن للمعلم القيمة المفروضة بالفرضية الصفرية فإن قيمـــته تختلف عنها بغض النظر عن كون هذه القيمة المقبولة أعلى أو أدنى من القيمة المفروضة بالفرضية الصفرية.

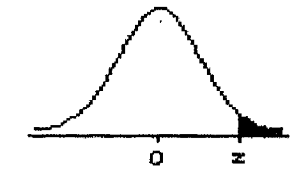
وبعـــبارة أخرى فإن الفرضية البديلة هي فرضية تقر بوجود الفروق دون أن تحدد إتجاه تلك الفروق.

تعني أن وسط المجتمع يختلف عن 70 $H_1: \mu \neq 70$

وسط المجتمع الأول يختلف عن وسط المجتمع الثاني $H_1: \mu 1 \neq \mu 2$

الفرضية البديلة المتجهة Directional Alternate Hypothesis





تهتم بكون المعلم اكبر أو اصغر من القيمة التي تفترضها الفرضية الصفرية

تعني أن وسط المجتمع اكبر من 70 $H_1: \mu > 70$

تعني أن وسط المجتمع اقل من 70 $H_1: \mu < 70$

وسط المجتمع الأول اكبر من وسط المجتمع الثاني $H_1: \mu 1 > \mu 2$

وعموماً فإن الاختبارات الاحصائية تستخدم البرهان غير المباشر للاستدلال واتخاذ القرار المتعلق بالفرضية البديلة من خلال فحص القرضية الصفرية التي تسمى ايضاً بالفرضية الاحصائية، أي أن الفرضية البديلة لا تخضع للاختبار الإحصائي والفرضية الصفرية هي التي تخضع للاختبار الإحصائي.

وحيث أن الفرضية تشير دائماً إلى عدم وجود فروق، فإن الفرضية البديلة هي التي تحدد كون الاختبار الإحصائي بذيل واحد أو بذيلين.

فإذا كانت الفرضية البديلة غير متجهة يكون الاختبار الإحصائي بذيلين.

وإذا كانت الفرضية البديلة متجهة يكون الاختبار الإحصائي بذيل واحد.

مثال: يدعي طالب في الثانوية العامة أن معدل عدد ساعات الدراسة اليومية له = 6 ساعات. فإذا اردنا اختبار هذا الإدعاء، فكيف تكتب كل من الفرضية الصفرية والفرضية البديلة.

$H_0: \mu = 6$
$H_1: \mu \neq 6$

2-4 اختبار الفرضيات الإحصائية
Testing Statistical Hypothesis

إن اتخـــاذ قـــرار حول ما إذا كانت الفرضية الصفرية مقبولة أم مرفوضة يتم عن طريق إختبار إحصائي.

الإختـــبار الإحصائي Statistical Test: هو متغير عشوائي ذو توزيع احتمالي يصف العلاقـــة بين القيم النظرية للمعلم والقيم المحسوبة من العينة. ويتم إتخاذ القرار المتعلق بقبول أو رفـــض الفرضـــية الصفرية، وبعد مقارنة قيمة الإختبار الإحصائي المحسوبة من العينة مع قيمته الحرجة المستخرجة من جداول خاصة لنتمكن من إتخاذ القرار.

وعلـــى أي حال فإن قبول الفرضية الصفرية لا يعني بالضرورة أنها صحيحة إذ انه يكون ناتجاً عن عدم وجود أدلة كافية من بيانات العينة لرفضها. ولهذا فإن التعبير الأصح في هذه الحالة هو القول ب " الفشل في رفض الفرضية الصفرية" وليس "قبول الفرضية الصفرية".

كما أن رفضها لا يعني أنها خاطئة بل يعني أن الإحصائي كان بعيداً عن المعلم المناظر له في المجتمع.

3-4 مفاهيم أساسية في فحص الفرضيات
Basic Concepts in Hypothesis Testing

لنتذكـــر أننا في فحص الفرضيات نقوم بتوظيف إختبار إحصائي لفحص فرضية صفرية (هـــي في حقيقة الأمر إما أن تكون صحيحة وإما أن تكون خاطئه) من اجل أن نصل الى قرار برفـــضها أو الفشل في رفضها (قبولها تجاوزاً). وعليه فإننا في هذا السياق نكون في واحدة من الحالات الأربعة التالية:

		الفرضية الصفرية	
		صحيحة	**خاطئه**
القرار	رفض	خطأ من النوع الأول	قرار صائب
	قبول	قرار صائب	خطأ من النوع الثاني

الخطأ من النوع الأول والخطأ من النوع الثاني Type One and Type Two Errors

يظهر نوعان من الخطأ عند اتخاذ القرار حول الفرضية الصفرية:

* **الخطـــأ مـــن النوع الأول Type One Error**: إتخاذ قرار برفض الفرضية الصفرية وهي في حقيقة الأمر صحيحة. وهذا قرار خاطيء نكون قد وقعنا فيه.
* **الخطأ من النوع الثاني Type Two Error**: إتخاذ قرار بقبول الفرضية الصفرية وهي في حقيقة الأمر خاطئة. وهذا قرار خاطيء نكون قد وقعنا فيه.
* إتخاذ قرار برفض الفرضية الصفرية وهي في حقيقة الأمر خاطئة. وهذا قرار صائب لا غبار عليه.
* إتخاذ قرار بقبول الفرضية الصفرية وهي في حقيقة الأمر صحيحة. وهذا قرار صائب لا غبار عليه.

مستوى الدلالة وقوة الاختبار Significant Level and Power of the Test

مستوى الدلالة(α) Significant Level : الحد الأعلى لاحتمال الوقوع في الخطأ من الـــنوع الأول، وهـــي تمثل مساحة الرفض تحت منحنى توزيع اختبار الاحصاءة، وهي احتمال رفض الفرضية الصفرية عندما تكون صحيحة، وتستخدم القيم 0.10 , 0.05 , 0.01 .

تكـــون القـــيمة القصوى لاحتمال ارتكاب الخطأ من النوع الأول (α) وهي مستوى الدلالـــة الإحـــصائية وقيمة مستوى الدلالة α يحددها الباحث لنفسه قبل جمع بياناته من عينة الدراسة.

فمثلاً $\alpha = 0.05$ تعني إذا تكررت التجربة لعدد كبير جداً فمن المحتمل أن نرفض فرضية صـــفرية وهـــي في الواقـــع صحيحة 5 مرات في كل 100 مرة، وان الاستنتاج يكون سليماً وصائباً بثقة 95%، تكون القيمة القصوى لاحتمال ارتكاب الخطأ من النوع الثاني (β).

هناك علاقة بين α ، β فزيادة احدهما يرافقها نقصان الآخر ولكن ليس بنفس المقدار.

* مستوى الثقة = $1 - \alpha$

* قوة الاختبار **Power of the Test (1 - β)**

قدرة الاختبار على رفض الفرضية الصفرية عندما تكون في حقيقة الأمر خاطئه.

$$\text{Power of the test} = 1 - \beta$$

يمكـــن أن نعـــين حـــداً حرجاً للرفض أو القبول للفرضية الصفرية مع الأخذ بالحسبان الفرضية البديلة.

كلما كبرت β يضعف الاختبار، وان تصغير β يحتاج لتكبير α.

اتخـــاذ قـــرارات لاحقـــة على أساس الدلالة الإحصائية قضية جدلية عند الإحصائيين، فالبعض يرى أن اختبارات الدلالة الإحصائية أفسدت الأبحاث العلمية ، لذلك لا بد من التوجه إلى الدلالة العلمية.

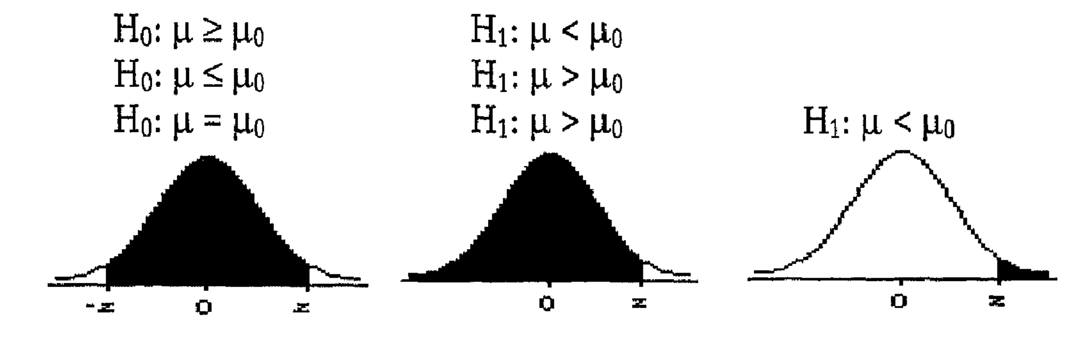
$$\beta = \sqrt{n}\,(\mu - \mu_0) / \sigma$$

* العوامل المؤثرة في قوة الاختبار

1. حجم العينة Sample Size
2. مستوى الدلالة Level of Significance
3. علاقة القيمة الحقيقية للمعلم بقيمته في الفرضية الصفرية.
4. كون الاختبار بذيل أو ذيلين One Tail or Two Tail Test

* اختبار بذيل واختبار بذيلين **One and Two Tails of Test**

الانحراف عن الفرضية الصفرية باتجاه واحد أو باتجاهين.



القرار بشأن الاختبار Decision Making

إذا كانـــت القيمة المحسوبة تقع في منطقة الرفض أي أنها أكبر من القيمة الجدولية نرفض الفرضية الصفرية H0 إذا كانـــت القـــيمة المحسوبة تقع في منطقة القبول أي أنها أقل من القيمة الجدولية نقبل الفرضية الصفرية H0

4-4 خطوات اختبار الفرضيات Hypothesis Testing Steps

1. تحديد نوع توزيع المجتمع.
2. صياغة الفرضيتين الصفرية والبديلة.
3. تحديد مستوى الدلالة (α) المناسب.
4. تحديد الاختبار الإحصائي المناسب لاختبار الفرضية الصفرية.
5. إذا كانت P اقل من α نرفض الفرضية الصفرية وبعكس ذلك نقبل هذه الفرضية.

* إذا كانت الفرضية البديلة غير متجهة $H_1 : \mu = C$

مـــستوى الدلالـــة يقسم إلى نصفين بالتساوي على كل من ذيلي توزيع المعاينة وتكون منطقة قبول الفرضية في الوسط ومنطقة رفضها على الذيلين.

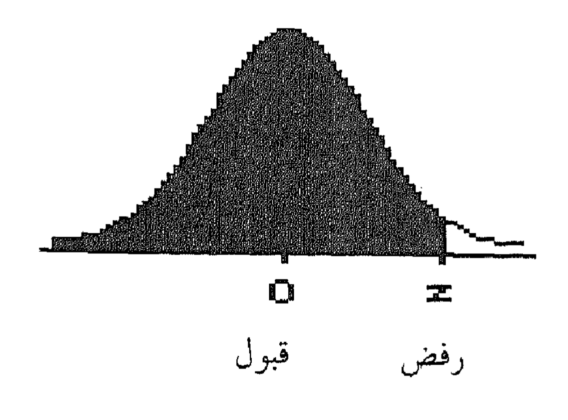


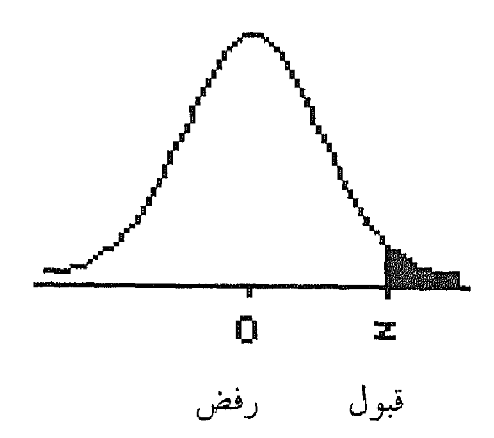
* إذا كانت الفرضية البديلة متجهة

إن منطقة الرفض ستكون إما في الذيل الأيمن أو الأيسر.

إن منطقة الرفض ستكون في الذيل الأيمن. $H_1 : \mu > C$

إن منطقة الرفض ستكون في الذيل الأيسر. $H_1 : \mu < C$



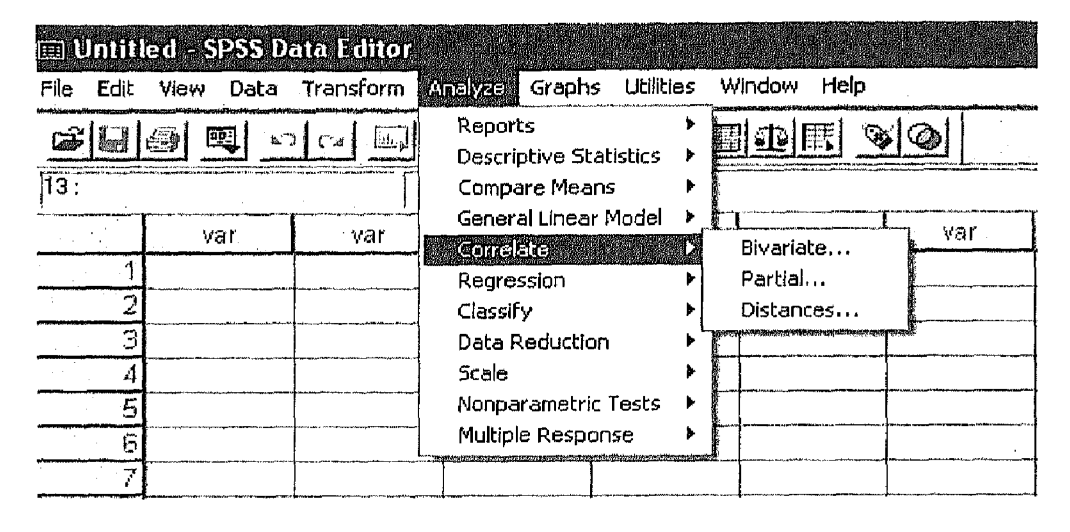


5-4 استخدام برمجية SPSS في حل المسائل.

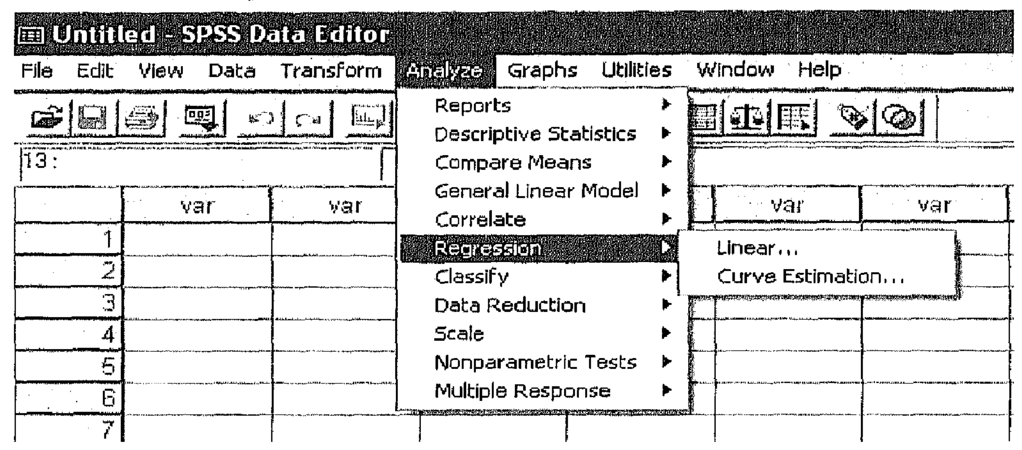
تستخدم برمجية SPSS لاختبار الفرضيات حول الاوساط كالتالي:

Untitled - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
Reports
Descriptive Statistics
Compare Means
General Linear Model
Correlate
Regression
Classify
Data Reduction
Scale
Nonparametric Tests
Multiple Response
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...

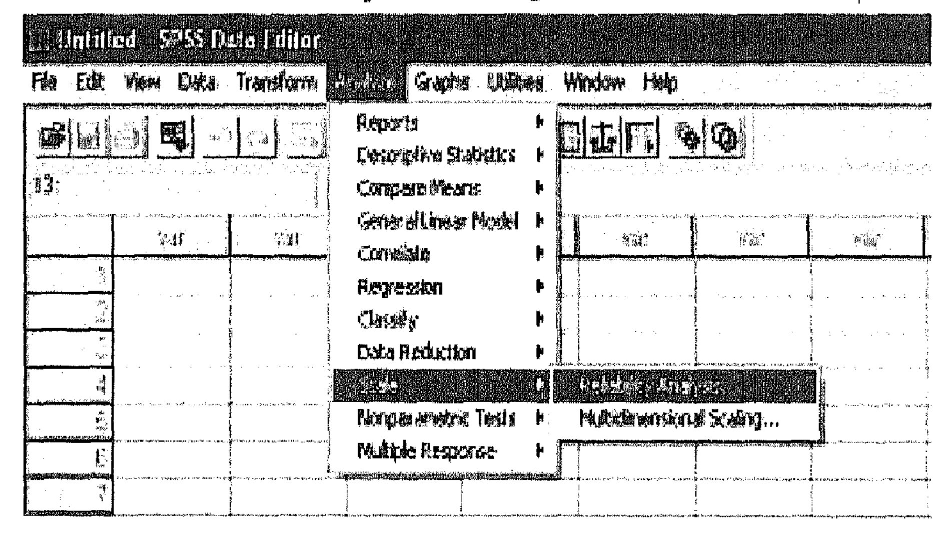
تستخدم برمجية SPSS لاختبار الفرضيات حول معاملات الارتباط كالتالي:



تستخدم برمجية SPSS لاختبار الفرضيات حول الانحدار كالتالي:



تستخدم برمجية SPSS لايجاد معامل الثبات كالتالي:



كما يمكن اختبار الفرضيات لأشياء كثيرة سنتعرف عليها في الفصول القادمة بشيء من التفصيل.

6-4 تمـاريـــن Exercise.

تمرين (1): ضع رمز الإجابة (نعم أو لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- يحقق اختبار الفرضيات والتقدير الإحصائي بفترة الغرض نفسه.

2- يستند اختبار الفرضيات على توزيعات المعاينة دائما.

3- تتمتع النتائج المنبثقة من اختبار الفرضيات الإحصائية بإمكانية تعميمها وأهميتها حيثما تم الوصول إليها.

4- تتمتع نتائج اختبار الفرضية العلمية بالعمومية والأهمية العلمية.

5- يتقرر كون الاختبار متجه أو غير متجه من الفرضية الصفرية.

6- تقل قوة الاختبار بازدياد حجم العينة.

7- الاختبار بذيلين أقل قوة من الاختبار بذيل واحد.

8- تصاغ الفرضيات الإحصائية حول المعالم فقط.

9- مستوى الدلالة هو الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الأول.

10- إذا كانت القيمة المحسوبة تقع في منطقة الرفض أي أنها أقل من القيمة الجدولية نقبل الفرضية الصفرية H_0

تمرين (2): ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- يتقرر كون الاختبار متجه أو غير متجه من:

أ- الفرضية الصفرية. ب- الفرضية البديلة. ج- مستوى الدلالة د- قوة الاختبار

2- يتم ارتكاب الخطأ من النوع الأول عند:

أ- قبول فرضية صفرية خاطئة.
ب- رفض فرضية صفرية صحيحة.
ج- رفض فرضية صفرية خاطئة.
د- قبول فرضية صفرية صحيحة.

3- يتم ارتكاب الخطأ من النوع الثاني عند:

أ- قبول فرضية صفرية خاطئة.
ب- رفض فرضية صفرية صحيحة.
ج- رفض فرضية صفرية خاطئة.
د- قبول فرضية صفرية صحيحة.

4- قوة الاختبار هي قدرة الاختبار على:

أ- قبول فرضية صفرية خاطئة.
ب- رفض فرضية صفرية صحيحة.
ج- رفض فرضية صفرية خاطئة.
د- قبول فرضية صفرية صحيحة.

5- تسمى القيمة العظمى لاحتمال ارتكاب الخطأ من النوع الأول ب:

أ- مستوى القبول.
ب- مستوى الرفض.
ج- مستوى الفصل.
د- مستوى الدلالة.

6- أي من التالية **ليست** من العوامل المؤثرة في قوة الاختبار:

أ- حجم العينة.
ب- كون الاختبار بذيل أو بذيلين.
ج- الإحصائي المستخدم.
د- مستوى الدلالة.

7- عندما يكون الاختبار بذيلين فإن قوة الاختبار تزداد:

أ- كلما اقتربت القيمة الحقيقية للمعلم مع القيمة المفروضة.
ب- كلما ابتعدت القيمة الحقيقية للمعلم عن القيمة المفروضة.
ب- كلما اقتربت القيمة الحقيقية للمعلم مع القيمة المفروضة.
ج- كلما تساوت القيمة الحقيقية للمعلم مع القيمة المفروضة.
د- لا شيء مما ذكر.

8- الذي يحدد كون الاختبار الإحصائي بذيل أم بذيلين هو:

أ- مستوى الثقة. ب- مستوى الدلالة. ج- الفرضية الصفرية. د- الفرضية البديلة.

9- مستوى الدلالة هو:

أ- الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الأول.

ب- الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الثاني.

ج- الحد الأدنى لاحتمال الوقوع في الخطأ من النوع الأول.

د- الحد الأدنى لاحتمال الوقوع في الخطأ من النوع الأول.

10- إذا كانت القيمة المحسوبة تقع في منطقة الرفض أي أنها أقل من القيمة الجدولية:

أ- نقبـــل الفرضية الصفرية H_0

ب- نقبـــل الفرضية البديـــلة H_1

ج- نرفض الفرضية الصفرية H_0

د- نرفض الفرضية البديلة H_1

الفصل الخامس

اختبار الفرضيات التي تتعلق بالمتوسطات الحسابية

Testing Hypothesis Regarding Mean

5-1 اختبار الفرضية المتعلقة بوسط حسابي واحد (مجتمع واحد)

* اختبار الفرضية المتعلقة بوسط حسابي واحد(حجم العينة كبير،تباين المجتمع معلوم)

Testing Hypothesis Regarding Mean (Big sample&σ Known)

* اختبار الفرضية المتعلقة بوسط حسابي واحد(حجم العينة كبير،تباين المجتمع غيرمعلوم)

Testing Hypothesis Regarding Mean (σ unknown, Big Sample)

* اختبار الفرضية المتعلقة بمتوسط حسابي واحد(العينة صغيرة، تباين المجتمع غير معلوم)

Testing Hypothesis Regarding Mean (σ unknown, Small Sample)

5-2 اختبار الفرضيات حول الفرق بين وسطين حسابيين

* اختبار الفرضيات حول الفرق بين وسطين للبيانات المستقلة
* اختبار الفرضيات حول الفرق بين وسطين للبيانات غير المستقلة

5-3 استخدام برمجية SPSS في حل المسائل.

5-4 تمارين Exercise.

الفصل الخامس

اختبار الفرضيات التي تتعلق بالمتوسطات الحسابية

Testing Hypothesis Regarding Mean

1-5 اختبار الفرضيات المتعلقة بوسط حسابي واحد (مجتمع واحد)

اختبار متوسط العينة 'X لمعرفة فيما إذا كان هناك فرق جوهري بين متوسط العينة ومتوسط المجتمع μ المأخوذة منه، أي هل العينة مسحوبة من المجتمع وممثلة له، ويمكن أن يستخدم الاختبار الإحصائي Z أو T.

ويستخدم الاختبار الإحصائي T لفحص فرضية تتعلق بالوسط الحسابي وهو اختبار One Sample T-Test ، ويجب أن يتحقق شرطان قبل إجراء الاختبار وهما:

1. يجب أن يكون المتغير موزعاً توزيعاً طبيعياً، ويستعاض عن هذا الشرط بزيادة حجم العينة، وتعتبر العينة كبيرة إذا كان حجمها 30 فأكثر.
2. أن تكون العينة عشوائية وقيم أفرادها لا تعتمد على بعضها البعض.

يستخدم هذا الاختبار لفحص ما إذا كان متوسط متغير ما لعينة واحدة يساوي قيمة ثابتة، وتكتب الفرضية المتعلقة بهذا الاختبار كالتالي:

$H_0: \mu = a$

حيث a قيمة ثابتة.

ويمكن تحديد قيمة الثابت a كما يلي:

1- العلامة الوسطى على تدريج ما.

- مقياس يتكون من 0 1 2 3 4 5 6 7 8 9 10

المتوسط = 5، لأنها تتوسط مدى الإجابة، يعني 11/55 = 5 ، وهي قيمة الثابت a

- مقياس ليكرت الخماسي يتكون من 1 2 3 4 5

المتوسط = 3، لأنها تتوسط مدى الإجابة، يعني 5/15 = 3 ، وهي قيمة الثابت a

الإجابات التي تقل عن قيمة الثابت a تعني فعالية متدنية (سالبة)، الإجابات التي تزيد عن قيمة الثابت a تعني فعالية عالية (موجبة).

2- من خلال المعلومات السابقة.

مقياس القلق المقنن وله متوسط = 50 درجه.

مقياس الذكاء وله وسط حسابي = 100

الاختبارات المقننة.

3- عدد الإجابات الصحيحة بطريقة الصدفة في امتحان ما.

اختبار مكون من 40 فقرة لكل منها 4 بدائل

احتمال إجابة أي سؤال بطريقة عشوائية = 0.25

عدد الإجابات المتوقع أن تكون صحيحة على جميع الاختبار=40*0.25 = 10 درجة

إذا قــل متوســـط الإجابات عن 10 يكون الاختبار صعب، وإذا زاد عنها فإننا نرفض فرضية الباحث وهي أن الاختبار ملائم.

1- اختبار الفرضية المتعلقة بوسط حسابي واحد(حجم العينة كبير،تباين المجتمع معلوم)

Testing Hypothesis Regarding Mean (Big sample&σ Known)

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة اكبر من 30 ، وتباين المجتمع معلوم، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{\sigma_x / \sqrt{n}}$$

مـــثال5-1) يعتقد مدير مدرسة الهمة الثانوية أن متوسط أداء طلابه في امتحان الثانوية العامـــة لا يختلف عن المتوسط العام لطلاب المملكة. فإذا علم أن عدد طلاب مدرسته يساوي $n= 81$ وأن متوسط معدلاتهم كان $X'= 58$ وأن متوسط المعدلات لطلاب المملكة كان 60 = μ بانحـــراف معياري يساوي $\sigma_x =8$ ، فهل تدعم هذه البيانات ما يعتقده المدير؟ استخدم $\alpha = 0.05$

*** خطوات اختبار الفرضيات**

1- تحديد نوع توزيع المجتمع.

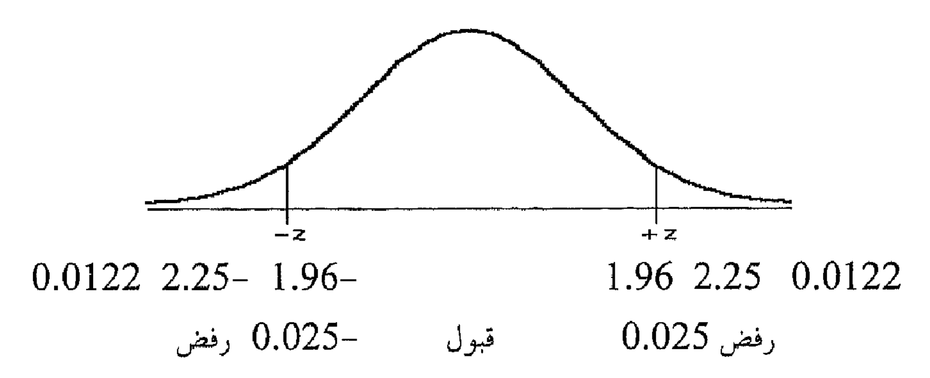
توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة اكبر من 30 ، لذلك نستعمل اختبار (Z).

2- صياغة الفرضيتين الصفرية والبديلة.

تعني أن وسط المجتمع لا يختلف عن 60 $H_0: \mu = 60$

تعني أن وسط المجتمع يختــلف عن 60 $H_1: \mu \neq 60$

3- تحديد مستوى الدلالة (α) المناسب. الفرضية البديلة، استخدم $\alpha = 0.05$



4- تحديد الاختبار الإحصائي المناسب لاختبار الفرضية الصفرية.

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة اكبر من 30 ، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{\frac{\sigma_x}{\sqrt{n}}} = \frac{58-60}{\frac{8}{\sqrt{81}}} = \frac{-2}{\frac{8}{9}} = -2.25$$

5- إذا كانــت (القــيمة المحسوبة) اكبر من > القيمة الجدولية α نرفض الفرضية الصفرية وبعكس ذلك نقبل الفرضية الصفرية.

إذا كانــت قيمة إحصاءة الاختبار ضمن منطقة الرفض يتم رفض الفرضية الصفرية H_0 وغير ذلك لا نستطيع رفضها ونقبل الفرضية البديلة H_1.

من الجدول: المساحة الواقعة دون -2.25 تساوي 0.0122

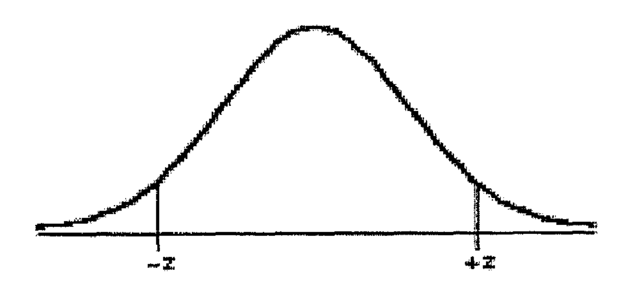
بما أن القيمة المحسوبة ($z=2.25$ 0.0122) > القيمة الجدولية ($z=.96$ 0.025 $=\alpha/2=0.05/2$) نرفض الفرضية الصفرية ونقبل البديلة، وهذا يعني أن أداء طلاب مدرسة الهمة في امتحان الثانوية العامة يختلف عن المتوسط العام لطلاب المملكة.

6- عن طريق الحدود الحرجة

الحدود الحرجة عند استخدام اختبار Z ومستوى الدلالة $\alpha = 0.05$ و $\alpha = 0.01$

نوع الاختبار/مستوى الدلالة	$\alpha = 0.05$	$\alpha = 0.01$
بذيلين	1.96±	2.58±
بذيل	1.64±	2.34±

بما أن $Z = -2.25$ وهي تقع في منطقة الرفض اذاً نرفض الفرضية الصفرية ونقبل الفرضية البديلة.

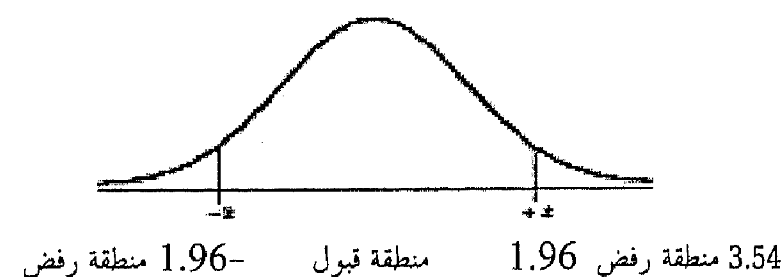


منطقة رفض 1.96 منطقة قبول -1.96 -2.25منطقة رفض

مثال5-2) إذا كان متوسط علامات الطلاب في مادة الإحصاء هو 70 بانحراف معياري مقداره 10 درجات وسحبت عينة عشوائية حجمها 50 طالباً وكان معدل علاماتهم 75، هل هناك فروق جوهرية بين العلامات بمستوى دلالة $\alpha = 0.05$ ؟

الحل: حجم العينة كبير وتباين المجتمع معلوم لذلك نستخدم احصاءة Z وهي:

$$Z = \frac{X' - \mu}{\sigma_x/\sqrt{n}} = \frac{75-70}{10/\sqrt{50}} = \frac{5}{10/7.07} = 3.54$$



منطقة رفض 1.96-   منطقة قبول   1.96 منطقة رفض 3.54

القرار: يمة Z المحســـــوبة = 3.54 قيمة $Z_{\alpha/2}$ الجدولية = 1.96

بما أن قيمة Z المحسوبة (3.54) > أكبر من قيمة $Z_{\alpha/2}$ الجدولية (1.96) وهي تقع في منطقة الرفض للفرضية الصفرية، لذلك نرفض H_0 وهذا يعني أن مستوى طلبة العينة يختلف عن المستوى العام بدرجة ثقة 95%.

2- اختبار الفرضية المتعلقة بوسط حسابي واحد(حجم العينة كبير،تباين المجتمع غيرمعلوم)

Testing Hypothesis Regarding Mean (σ unknown, Big Sample)

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة اكبر من 30 ، وتباين المجتمع غير معلوم، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{S_x/\sqrt{n}}$$

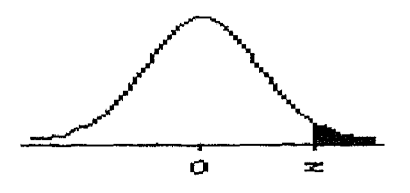
مــثال(5-3) إذا كانـــت أوزان الطلاب في المدرسة تخضع لتوزيع طبيعي وسطه 60 كغم، وسحبت منه عينة عشوائية حجمها 50 طالباً وسجلت متوسطاً قدره 65كغم بانحراف معياري قدره 10كغم، اختبر الفرضية ان المتوسط اكبر من 60 باستخدام $\alpha = 0.05$ ؟

$H_0: \mu = 60$ تعني أن وسط المجتمع لا يختلف عن 60

$H_1: \mu > 60$ تعني أن وسط المجتمع اكـــــبر من 60

$$Z = \frac{X' - \mu}{S_x/\sqrt{n}} = \frac{65-60}{10/\sqrt{50}} = \frac{5}{10/7.07} = 3.54$$

بمـــا أن الاختــبار متجه أي بذيل واحد فإن مستوى الدلالة هو $\alpha = 0.05$ والقيمة الجدولية $Z\alpha = 1.645$



منطقة الرفض 1.64 منطقة عدم الرفض

القرار:

بمـــا أن قيمة Z المحسوبة (3.54) > أكبر من قيمة Z_α الجدولية (1.64) وهي تقع في منطقة الرفض لذلك نرفض H_0 ونقبل الفرضية البديلة وهذا يعني أن معدل وزن العينة اكبر من 60 وهو المستوى العام بدرجة ثقة 95%.

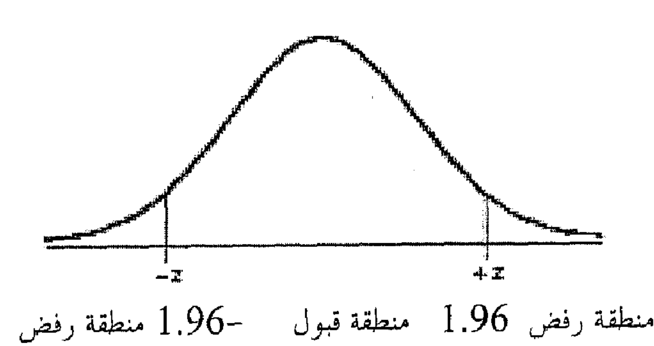
* اذا كانت الفرضية غير متجهة

تعني أن وسط المجتمع لا يختلف عن 60 $H_0: \mu = 60$

تعني أن وسط المجتمع يختلف عن 60 $H_1: \mu \neq 60$

$$Z = \frac{X' - \mu}{S_x/\sqrt{n}} = \frac{65-60}{10/\sqrt{50}} = \frac{5}{10/7.07} = 3.54$$

بمـــا أن الاختـــبار غـــير مـــتجه أي بـــذيلين فإن مستوى الدلالة هو 2/ $\alpha = 0.05$ = 0.025 والقيمة الجدولية $Z\alpha/2 = 1.96$



منطقة رفض 1.96 منطقة قبول -1.96 منطقة رفض

القرار:

بمـــا أن قيمة Z المحسوبة (3.54) > أكبر من قيمة Z_α الجدولية (1.96) وهي تقع في مـــنطقة الرفض لذلك نرفض H_0 ونقبل الفرضية البديلة وهذا يعني أن معدل وزن العينة يختلف عن 60.

3- اختبار الفرضية المتعلقة بمتوسط حسابي واحد(σ غير معلومة والعينة صغيرة الحجم)

Testing Hypothesis Regarding Mean (σ unknown, Small Sample)

عـــندما يكون الانحراف المعياري لأداء المجتمع σ غير معلوم، والعينة صغيرة الحجم، فإن الاختـــبار المناسب هو اختبار T، وفي هذه الحالة نقدر الخطأ المعياري لتوزيع المعاينة من خلال العلاقة

Sx انحراف معياري عينة

$$\sigma x' = Sx' = \frac{Sx}{\sqrt{n}}$$

$$T = \frac{X' - \mu}{Sx / \sqrt{n}}$$

تتقرر الحدود الحرجة للرفض أو القبول في ضوء درجات الحرية γ ومستوى الدلالة α ، وكون الفرضية البديلة متجهة أو غير متجهة.

مثال5-4)

1- استخدم برنامج لتعليم اللغة الإنجليزية مع 16 طالباً من طلاب جامعة الإسراء، وقد اعـــتقد مصممو البرنامج أنه يرفع متوسط أداء الطلاب إلى أكثر من 0.70 وعند إنهاء الطلاب للبرنامج أجري اختبار مقنن فكان متوسط أدائهم = 72.6 بانحراف معياري غير متحيز = 5 ، هل تدعم هذه النتائج صحة ادعاء مصممي البرنامج؟ استخدم $\alpha = 0.05$

2- وإذا كـــان ادعاء مصممي البرنامج أن برنامجهم يجعل علامات الطلاب الضعاف لا تخـــتلف عـــن المتوســـط العام وهو 70 ، فهل تدعم البيانات صحة ادعاء مصممي البرنامج؟ استخدم $\alpha = 0.05$

1- حل الحالة الأولى:

H_0: $\mu > 70$
H_1: $\mu <= 70$

عندما يكون الانحراف المعياري لأداء المجتمع σ غير معلوم، فإن الاختبار المناسبT

نحدد القيم الحرجة للرفض أو القبول:

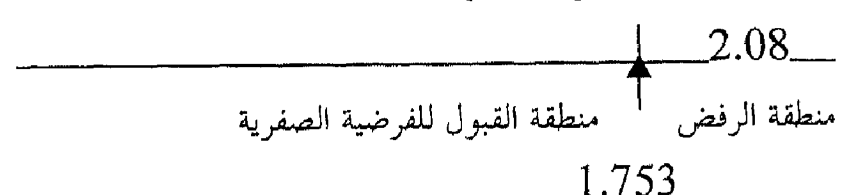
بما أن مستوى الدلالة $\alpha = 0.05$

درجة الحرية $\gamma = n-1$ ، $= 16 - 1 = 15$

الاختبار بذيل واحد

بالاستعانة بالجدول رقم 3 تكون القيمة الحرجة $T_{15,0..95} = 1.753$

تكون منطقتا القبول والرفض كما هو مبين أدناه



الآن نحسب قيمة T كما يلي

$$T = \frac{X' - \mu}{\frac{Sx}{\sqrt{n}}} = \frac{72.6 - 70}{\frac{5}{\sqrt{16}}} = 2.08$$

بما أن قيمة T المحسوبة (2.08) تقع في منطقة الرفض للفرضية الصفرية، لذلك نرفض الفرضية الصفرية ونقبل البديله ونقول أن متوسط أداء الطلاب الذين خضعوا للبرنامج لم يكن أعلى من 70 بمستوى الدلالة $\alpha = 0.05$

2- وإذا كان ادعاء مصممي البرنامج أن برنامجهم يجعل علامات الطلاب الضعاف لا تختلف عن المتوسط العام وهو 70 ، فهل تدعم البيانات صحة ادعاء مصممي البرنامج؟ استخدم $\alpha = 0.05$

$H_0: \mu = 70$
$H_1: \mu \neq 70$

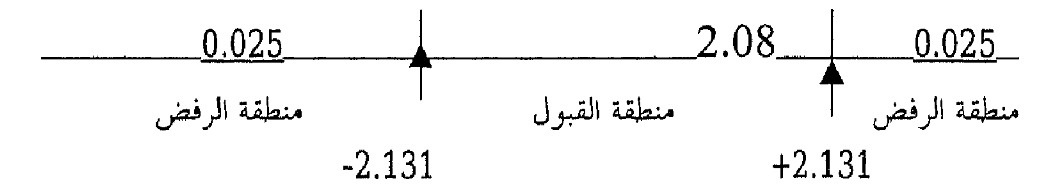
بما أن مستوى الدلالة $\alpha = 0.05$ والاختبار بذيلين

درجة الحرية $= n-1 = 1-16 = 15$

درجة الحرية $\gamma = n-1$ ، $= 16 - 1 = 15$

بالاستعانة بجدول توزيع t تكون القيمة الحرجة $T_{15,0.95} = \pm 2.131$

تكون منطقتا القبول والرفض كما هو مبين أدناه



الآن نحسب قيمة T كما يلي

$$T = \frac{X' - \mu}{\frac{Sx}{\sqrt{n}}} = \frac{70 - 72.6}{\frac{5}{\sqrt{16}}} = 2.08$$

بما أن قيمة T المحسوبة (2.08) تقع في منطقة القبول للفرضية الصفرية، لذلك نقبل الفرضية الصفرية ونقول أن متوسط أداء الطلاب الذين خضعوا للبرنامج لم يختلف عن المتوسط العام وهو 70 بمستوى الدلالة $\alpha = 0.05$

2-5 اختبار الفرضيات حول الفرق بين وسطين حسابيين Testing Hypothesis Regarding the Difference between Two Means

تستخدم لإجراء المقارنة بين وسطين لمجتمعين في حالة البيانات المستقلة والبيانات غير المستقلة.

1- اختبار الفرضيات حول الفرق بين وسطين حسابيين (للبيانات المستقلة).

البيانات المستقلة: البيانات التي لا يوجد فيما بينها ارتباط. أداء مجموعة للذكور وأخرى من الإناث.

X'_1 وسط العينة الأولى مستقلاً عن X'_2 وسط العينة الثانية، الفرضية غير متجهة

$H0: \mu_1 - \mu_2 = 0$
$H1: \mu_1 - \mu_2 \neq 0$

الفرضية عندما تكون متجهة

$H0: \mu_1 - \mu_2 > 0$
$H1: \mu_1 - \mu_2 < 0$

يتخذ التوزيع شكل T بدرجات حرية $\gamma = n1 + n2 - 2$

n1: حجم العينة الأولى.

n2: حجم العينة الثانية.

الاختبار الإحصائي المناسب:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_d \sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

الافتراضات التي يجب توفرها لتطبيق اختبار T

1. التوزيع الطبيعي Normality

يجــب أن يكــون توزيع متغير الاختبار طبيعياً في كل فئة من فئات متغير التجميع، واذا كانت العينة 30 فأكثر يمكن الاستغناء عن هذا الشرط.

2. تجانس التباين في المجتمعين Homogeneity

يجب أن يكون تباين متغير الاختبار متساوياً في كلا فئتي متغير التجميع.

3. العشوائية، والاستقلالية Independence

مــثال5-5) أراد باحــثاً أن يعرف فعالية اسلوب معين في التدريس لطلبة الجامعة في مــستوى السنة الأولى، فأخذ شعبتين كل شعبة تتكون من 25 طالب عشوائياً في مجموعتين ثم عــين عــشوائياً أحــداهما لتكون مجموعة تجريبية والأخرى ضابطة.وفي نهاية التجربة أعطيت المجموعتين اختباراً موحداً فكانت النتائج كما يلي

المجموعة التجريبية	المجموعة الضابطة
$n1 = 25$	$n2 = 25$
$X'_1 = 7.65$	$X'_2 = 6$
$Sx_1 = 2.55$	$Sx_2 = 2.43$

هــل تــدل هذه البيانات على أن أداء المجموعة التجريبية كان أفضل من أداء المجموعة الضابطة على مستوى دلالة $\alpha = 0.05$

الفرضية الصفرية $H0: \mu_1 = \mu_2$

الفرضية عندما تكون متجهة $H1: \mu_1 > \mu_2$

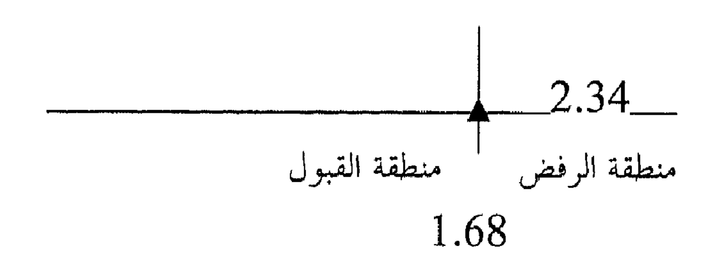
نحدد القيمة الحرجة للرفض أو القبول:

بما أن مستوى الدلالة $\alpha = 0.05$ ، والاختبار بذيل واحد.

ودرجات الحرية $\gamma = 25 + 25 - 2 = 48$

تكون قيمة T الحرجة من الجدول $= 1.68 = T_{48,0.05}$

$\alpha = 0.05$



نحسب الخطأ المعياري لتوزيع المعاينة كما يلي:

$$S^2_d = \frac{(n_1 - 1)\, S^2_1 + (n_2 - 1)\, S^2_2}{n_1 + n_2 - 2} = \frac{24 * 2.55^2 + 24 * 2.43^2}{25 + 25 - 2} = 6.204$$

$S_d = 2.491$

نحسب قيمة الإحصائي T كما يلي:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$T = \frac{X'_1 - X'_2}{S_d \sqrt{1/n_1 + 1/n_2}} = \frac{7.65 - 6}{2.491 \sqrt{1/25 + 1/25}} = 2.34$$

نقارن القيمة المحسوبة (2.34) بالقيمة الحرجة (1.68) فنجد أنها تقع في منطقة الرفض.

إن البيانات تدل على أن الذين يخضعون للبرنامج التدريبي يصبح أداؤهم أفضل من الذين لا يخضعون على مستوى دلالة $\alpha = 0.05$

تبعات عدم الإيفاء بافتراضات اختبار T

1- التوزيع الطبيعي Normality

2- تجانس التباين في المجتمعين Homogeneity

3- الاستقلالية Independence

1- افتراض التوزيع الطبيعي Normality

إن هذا الافتراض يمكن مخالفته بدون تبعات تذكر.

2- افتراض الاستقلالية Independence

يهدد البحث بأكمله وعلى الباحث أن لا يتهاون في ذلك اطلاقاً.

3- تجانس التباين في المجتمعين Homogeneity

■ يمكن مخالفته إذا تساوت العينتان في أعداد افر ادهما. إذا كانت n1=n2

■ ولكن عندما تكون n1≠ n2 فهناك وضعين:

1- العينة الكبيرة الحجم منتمية للمجتمع ذي التباين الكبير

والعينة الصغيرة الحجم منتمية للمجتمع ذي التباين الصغير

فـإن احـتمال ارتكاب الخطأ من النوع الأول تكون أقل من α إذاً يكون الباحث في الجانب الأمين.

2- العينة الكبيرة الحجم منتمية للمجتمع ذي التباين الصغير

والعينة الصغيرة الحجم منتمية للمجتمع ذي التباين الكبير

فإن احتمال ارتكاب الخطأ من النوع الأول تزداد.

2- اختبار الفرضيات حول الفرق بين وسطين حسابيين (للبيانات غير المستقلة).

البيانات غير المستقلة: البيانات التي يوجد فيما بينها ارتباط.

مثال: إجراء اختبار على نفس المجموعة مرتين، n1 = n2 = n

مثال: عينة من الأزواج وزوجاتهم.

X'_1 وسط العينة الأولى مرتبطاً مع X'_2 وسط العينة الثانية وهي نفس العينة الأولى.

اختبار T للعينات المترابطة Paired Sample T-Test: يستخدم لفحص فرضية متعلقة بمساواة متوسط متغيرين أو مساواة متوسط متغير لعينتين غير مستقلتين.

لضمان دقة نتائج اختبار T يجب تحقق الشروط التالية:

1. أن يكون توزيع الفرق بين المتغيرين طبيعياً. ويمكن تجاهله إذا كان حجم العينة أكثر من 30

2. أن تكون العينة عشوائية.

الفرضية عندما تكون غير متجهة

$H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

الفرضية عندما تكون متجهة

$H_0: \mu_1 - \mu_2 > 0$
$H_1: \mu_1 - \mu_2 < 0$

يتخذ التوزيع شكل T بدرجات حرية $\gamma = n - 1$

n: عدد أزواج المشاهدات. وسط التوزيع =0

الانحراف المعياري، الخطأ المعياري لتوزيع المعاينة لفروق الأوساط $Sx'_1\text{-}x'_2$

$$Sx'_1\text{-}x'_2 = \sqrt{S^2 x'_1 + S^2 x'_2 - 2d \; S x'_1 \, S x'_2}$$

$$S^2 x'_1 = \frac{S^2 1}{n}$$

$$S^2 x'_2 = \frac{S^2 2}{n}$$

$$s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}}$$

مثال5-6) إذا كانت علامات طلاب شعبة في مادة الإحصاء في الاختبار الأول والاختبار الثاني كما هو مبين أدناه، والمطلوب اختبار الفرق بين متوسطي درجات الاختبارين في مادة الإحصاء.

الفرضية:

$H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

1- نحسب معدل الفروق ما بين العينتين:

$$d_i = X_i - Y_i$$

$$d' = \frac{\sum d_i}{n}$$

2- نحسب تباين الفروق:

$$S_d^2 = \frac{\sum d_i^2 - \frac{(\sum d')^2}{n}}{n - 1}.$$

3- نستخدم الاختبار t كما يلي:

$$T = \frac{\bar{X} - \mu_a}{\frac{S_a}{\sqrt{n}}}$$

Stud	Stat1_m	Stat2_m	di=Stat1-Stat2	di^2
1	20	15	5	25
2	20	14	6	36
3	18	20	-2	4
4	15	17	-2	4
5	16	17	-1	1
6	12	18	-6	36
7	15	20	-5	25
8	18	18	0	0
9	18	14	4	16
10	18	17	1	1
11	14	8	6	36
12	16	18	-2	4
13	12	14	-2	4
14	17	14	3	9
15	14	12	2	4
16	19	20	-1	1
17	11	10	1	1
18	10	10	0	0
19	19	20	-1	1
20	12	15	-3	9
21	17	20	-3	9
المجموع			0	226
d'			0	

$$S_d^2 = \frac{d_i^2 - \frac{(d')^2}{n}}{n - 1} = 11.3$$

$$S_d = 3.361547$$

$$T = \frac{\bar{X} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

$T = (0-0)/(3.36/4.58) = 0$ المحسوبة

$\gamma = n - 1 = 21-1 = 20$

$T_{20,0.05} = 1.72$ الجدولية

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات الاختبار الأول والثاني؟

بما أن القيمة المحسوبة (0) < أقل من القيمة الجدولية (1.72) وهي تقع في منطقة القبول، اذاً نقبل الفرضية الصفرية، ويتبين انه ليس هناك فرقاً بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني.

مثال5-7) أراد باحث مقارنة الرواتب التي يستلمها مجموعة من الموظفين من الأزواج والزوجات وكانت كما يلي:

الأزواج	الزوج	الزوجة	Di=Stat1-Stat2	di^2
1	240	230	10	100
2	260	270	-10	100
3	250	240	10	100
4	230	230	0	0
5	280	270	10	100
6	220	230	-10	100
7	240	250	-10	100
8	260	250	10	100
9	240	230	10	100
10	250	220	30	900
المجموع			50	1700
d'			5	

$$S_d^2 = \frac{d_i^2 - \frac{(d')^2}{n}}{n-1} = \frac{1697.5}{9} = 188.6111$$

$$S_d = 13.73358$$

$$T = \frac{\bar{X} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

$T = (5\text{-}0)/(13.73/3.16) = 5/4.43 = 1.13$ المحسوبة

$\gamma = n - 1 = 10 - 1 = 9$

$T_{9,0.05} = 1.83$ الجدولية

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات دخل الأزواج والزوجات ؟

بما أن القيمة المحسوبة (1.13) > أقل من القيمة الجدولية (1.83) وهي تقع في منطقة القبول، اذاً نقبل الفرضية الصفرية، ويتبين انه ليس هناك فرقاً بين دخل الأزواج والزوجات.

أهم الأساليب الإحصائية الشائعة واستخداماتها:

المقاييس البارامترية		المقاييس اللابارامترية	
اختبار (z) Z-test اختبار (ت) t-test	- تحديد ما إذا كان متوسطان أو نسبتان، أو معاملا ارتباط يختلفان عن بعضهما. - تحديد ما إذا كان متوسط واحد أو نسبة واحدة أو معامل ارتباط واحد يختلف عن تلك العلاقة للمجتمع.	كاي	- يستخدم لتقدير ما إذا كان توزيعان تكراريان تختلف عن بعضها بشكل دال.
تحليل التباين Analysis of variance One way Anova Two way Anova	- تحديد ما إذا كانت درجات المتوسط في عنصر أو اكثر تختلف عن بعضها. - ما إذا كان هناك تفاعل دال بين العناصر المختلفة. - يقيس إذا ما إذا كانت التباينات Variances مختلفة عن بعضها.	Mann-Whitney U test	يستخدم لقياس ما إذا كان متوسطين غير مرتبطين Uncorrected Means يختلفان بشكل دال

اختبارات تستخدم بعد تحليل التباين. Duncan's Scheffe's Tuky	تستخدم إذا ظهرت قيمة F دالة وذلك بهدف اختبار الدلالة الإحصائية للفروق بين متوسطات مجموعات محددة	Wilxoxon signed test	يستخدم لقياس ما إذا كان متوسطين مرتبطين Correlated Means يختلفان بشكل دال
تحليل التغاير Analysis of Covariance (Anacova)	مشابه في الاستخدام لأسلوب تحليل التباين إلا انه يمكن من ضبط متغير مستقل أو اكثر في المتغير التابع.	Kruskal-Wallis test	يستخدم لتقدير ما إذا كان 3 قيم أو اكثر للمتوسطات في عنصر واحد تختلف بدلالة إحصائية
Trend Analysis	لاختبار الاتجاه المفترض		
Confidence limits	يستخدم لتقدير قيمة في المجتمع بالاعتماد على القيمة المعروفة للعينة.		

الأساليب الإحصائية لحساب الفروق:

	عدد المتغيرات المستقلة	مستويات القياس (القياس هنا للمتغير التابع)		
		اسمي	رتبي	فئوي أو نسبي
عينة واحدة	1	كاي تربيع لحسن المطابقة	Smirnov	t-test للعينة الواحدة
عينتان مستقلتان	1	كاي تربيع للارتباط للعينات المستقلة ؛ Fisher exact test فشر	Man Whitney U-Test. مان وتني ؛ Median Test اختبار الوسيط	t-test للعينات المستقلة.
عينتان غير مستقلتان	1	كاي تربيع لنسبتين بيانات غير مستقلة	Wilcoxon Signed Rank Test. Sign Test. اختبار الإشارة	t-test للعينات المستقلة. Randomization Test. Walsh Test.
اكثر من عينتين مستقلتين	1	كاي للعينات المستقلة	Kruskal Walliss كروسكال والبس Median test	Analysis of Variance (One Way Anova). تحليلي التباين أحادى الاتجاه.

اكثر من عينتين غير مستقلتين	1	Chochran O test Friedman Test (البيانات الثنائية).	Friedman Test اختبار فريدمان	Anova (one way) اختبارات تستخدم بعد تحليل التباين. Duncan's Multiple-range, Scheffe's, test., Tuky
عينتان أو اكثر	2 أو اكثر	كاي تربيع		Factor analysis التحليل العاملي 2 way Anova تحليل التباين ثنائي الاتجاه. : Ancova

3-5 استخدام برمجية SPSS في حل المسائل.

1- اختبار الفرضية المتعلقة بوسط حسابي واحد.

اختبار T للعينة الواحدة One Sample T-Test

إذا كانت العلامات أدناه هي علامات الطلاب في مادة قواعد البيانات استخدم اختبار T لفحص وجود فرق بين متوسط درجة اختبار مادة قواعد البيانات (15.9286) وبين المتوسط الطبيعي وهو (12.5).

يمكن صياغة سؤال الدراسة بأحد الأشكال التالية:

هل هناك فرق بين متوسط درجة الطلاب وبين المتوسط الطبيعي؟

نرفض الفرضية إذا كانت دلالة قيمة Sig.(2-tailed) اكبر من المستوى المقبول (0.05) وهذا يعني أن المتوسط لا يساوي القيمة الثابتة a.

* ادخل البيانات كما هو مبين أدناه:

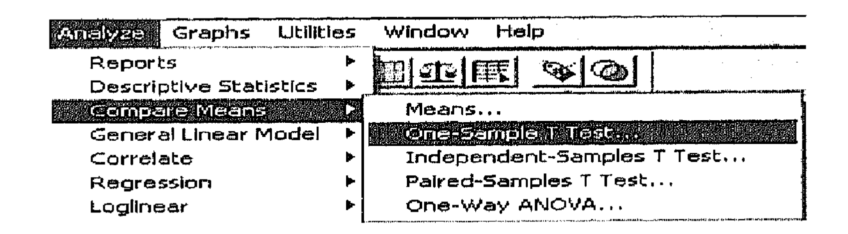
Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

14 :

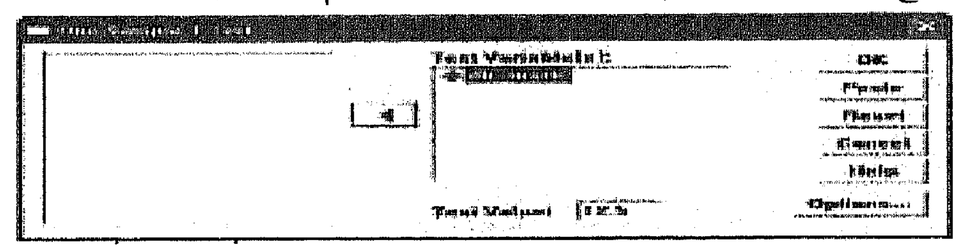
	db_mark	var	var	var	var
1	14.00				
2	14.00				
3	20.00				
4	17.00				
5	17.00				
6	18.00				
7	20.00				
8	18.00				
9	14.00				
10	17.00				
11	8.00				
12	18.00				
13	14.00				
14	14.00				
15					

Analyze – Compare Means – One-Sample T Test…



ضع المتغير db_mark في خانة: Test Variable(s):

وضع القيمة 12.5 في خانة: Test Value: - ثم أضغط زر OK



تظهر النتيجة التالية:

T-Test

One-Sample Statistics

	عدد المشاهدات N	Mean الوسط الحسابي	Std. Deviation الانحراف المعياري	Std. Error Mean الخطأ المعياري
DB_MARK	14	15.9286	3.1736	.8482

One-Sample Test

	Test Value = 12.5					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
DB_MARK	4.042	13	.001	3.4286	1.5962	5.2609

النتيجة:

تم حساب متوسط الفرق بين المتغير والقيمة المفترضة (Mean Difference) = 3.4286 والذي يشير إلى أن الوسط الحسابي لعينة الطلاب (15.9286) كان أعلى من المستوى الطبيعي (12.5)، وكان الفرق = 3.4286 وهل هذا الفرق كافي لأن نقرر أنه ذو دلالة إحصائية؟

$H_0: \mu = 12.5$

$H_1: \mu \neq 12.5$

بمـــا أن مـــستوى المعـــنوية Sig.(2 -tailed)=0.001 وهي أقل من < مستوى الدلالة α=0.05، اذاً نقبل الفرضية البديلة ونرفض الفرضية الصفرية

اســـتخدم اختبار T لفحص وجود فرق بين متوسط درجة اختبار مادة قواعد البيانات وبـــين المتوســـط الطبيعي (12.5) ، وقد تبين من خلال النتائج الموضحة في الجدول أعلاه أن متوســـط علامـــات الطلبة كان أعلى من المتوسط الطبيعي (12.5) فقد بلغ متوسط علامات الطـــلاب (15.93) بإنحـــراف معـــياري (3.17) وقد بلغت قيمة t=4.042 وهي ذات دلالة احصائية عند مستوى أقل من 0.05

كانت قيمة t=4.042 وبلغ مستوى دلالتها Sig.(2-tailed)=0.001 قيمة صغيرة جداً وهـــي أقل من المستوى المقبول لدينا (α=0.05) وهذا يعني ان متوسط الاختبار في مادة قواعد البيانات ≠12.5 بل هو أعلى من ذلك.

القيمة المحسوبة t = 4.042

القيمة الحرجة [+2.160   -2.160]

إن القيمة المحسوبة تقع في منطقة رفض الفرضية الصفرية يعني قبول البديلة.

المطلـــوب: إذا كانت العلامات أدناه هي علامات الطلاب في مادة الإحصاء استخدم اختـــبار T لفحص وجود فرق بين متوسط درجة اختبار مادة الإحصاء (15.7619) وبين المتوسط الطبيعي وهو (12.5).

$H_0: \mu = 12.5$

$H_1: \mu \neq 12.5$

	stat1_m
1	20.00
2	20.00
3	18.00
4	15.00
5	16.00
6	12.00
7	15.00
8	18.00
9	18.00
10	18.00
11	14.00
12	16.00
13	12.00
14	17.00
15	14.00
16	19.00
17	11.00
18	10.00
19	19.00
20	12.00
21	17.00

Data View

النتيجة: بما أن مستوى المعنوية Sig.(2-tailed)=0.000 ومستوى الدلالة α=0.05، إذا نقبل الفرضية البديلة ونرفض الفرضية الصفرية

يعني أن متوسط الاختبار في مادة الإحصاء ≠12.5 بل هو أعلى من ذلك.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
STAT_MAR	21	15.7619	3.0480	.6651

One-Sample Test

	Test Value = 12.5					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
STAT_MAR	4.904	20	.000	3.2619	1.8745	4.6494

2- اختبار الفرضيات حول الـــفـــرق بين وسطين حسابيين (للبيانات المترابطة)

اختبار T للعينات المترابطة Paired Sample T-Test

السؤال هو: هل تتساوى متوسطات الطلاب بالاختبارين؟

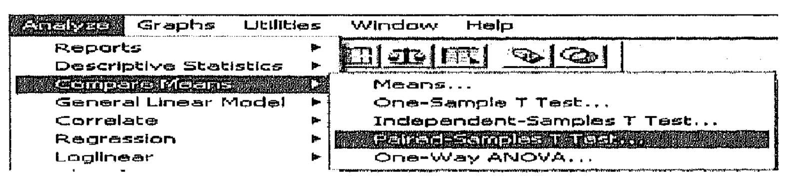
نرفض الفرضية اذا كانت دلالة قيمة Sig.(2-tailed) أقل من المستوى المقبول (0.05) وهذا يعني أن المتوسطين غير متساويين.

* ادخل البيانات كما هو مبين أدناه:

	stat1_m	stat2_m
1	20.00	15.00
2	20.00	14.00
3	18.00	20.00
4	15.00	17.00
5	16.00	17.00
6	12.00	18.00
7	15.00	20.00
8	18.00	18.00
9	18.00	14.00
10	18.00	17.00
11	14.00	8.00
12	16.00	18.00
13	12.00	14.00
14	17.00	14.00
15	14.00	12.00
16	19.00	20.00
17	11.00	10.00
18	10.00	10.00
19	19.00	20.00
20	12.00	15.00
21	17.00	20.00

Data View / Variable View

Analyze – Compare Means – Paired-Samples T Test…



ضع المتغيرات stat1_m و stat2_m في خانة Paired Variables:

ثم أضغط زر OK

تظهر لدينا المخرجات التالية:

المطلوب: إذا كانت العلامات أدناه هي علامات الطلاب في مادة الإحصاء في الاختبار الأول والاختـــبار الثاني لفحص هل لهما نفس المتوسط الحسابي، استخدم اختبار لفحص وجود فرق بين متوسطي درجات الاختبارين في مادة الإحصاء.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$ **T-Test**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	STAT1_M	15.7619	21	3.0480	.6651
	STAT2_M	15.7619	21	3.6319	.7925

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	STAT1_M & STAT2_M	21	.505	.020

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	STAT1_M - STAT2_M	.0000	3.3615	.7335	-1.5302	1.5302	.000	20	1.000

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات الاختبار الأول والثاني؟

بما أن مستوى المعنوية Sig.(2-tailed)=1.000 ومستوى الدلالة α=0.05، إذا نرفض الفرضية البديلة ونقبل الفرضية الصفرية.

تـــبين انه ليس هناك فرقاً بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني فقـــد بلغـــت قيمة (t=(0.000 وهي غير دالة إحصائيا على مستوى α=0.05 حيث حسب الـــبرنامج متوسط الفرق بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني والذي بلغ (0.0000).

3- اختبار الفرضيات حول الـــفـــرق بين وسطين حسابيين (للبيانات المستقلة)

اختبار T للعينات المستقلة Independent-Sample T-Test

فحـــص فرضية متعلقة بمساواة متوسط متغير ما لعينتين مستقلتين، وله شكلان الأول في حالة افتراض تساوي تباين العينتين، والثاني في حالة افتراض أن تباين العينتين غير متساوي.

إذا كانت علامات الطلاب والطالبات في مادة الإحصاء في شعبة ما كما هو مبين أدناه، فهل يختلف تحصيل الذكور عن الإناث في هذه الشعبة؟

هل يرتبط تحصيل الطلبة حسب الجنس؟

الفرضية المطلوب اختبارها:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

للاســـتخدام الاختبار يكون هناك متغير التجميع (Grouping Variable) وهو المتغير الذي يقسم العينة الكلية إلى عينتين جزئيتين غير متداخلتين، مثل متغير الجنس الذي يقسم العينة

الكلـــية إلى عينة ذكور وعينة إناث، والمتغير الثاني يسمى متغير الاختبار (Test Variable) أو المتغير التابع، وهو متغير كمي .

والهـــدف مـــن هـــذا الاختبار هو معرفة ما إذا كان متوسط متغير الاختبار لفئة متغير التجمـــيع الأولى (الذكور) مساوية لمتوسط متغير الاختبار لدى الفئة الثانية (الإناث) من متغير التجميع.

السؤال هو: هل يختلف تحصيل الطلاب في مادة الإحصاء باختلاف جنسهم؟

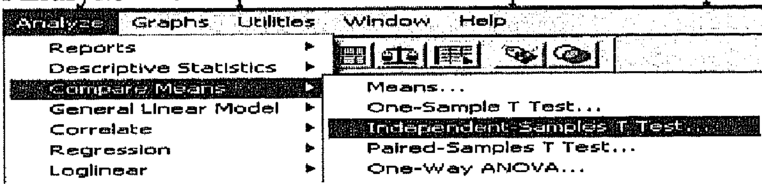
أو هل يرتبط تحصيل الطلاب في مادة الإحصاء بالجنس؟

نـــرفض الفرضية الصفرية إذا كانت دلالة قيمة مستوى الدلالة Sig.(2-tailed) أقل من المـــستوى المقـــبول (0.05) وهذا يعني أن المتوسطين غير متساويين، وذلك بعد تحديد قيمة t المستخدمة بناء على نتيجة اختبار Levene Test لمساواة تباين العينتين.

* ادخل البيانات كما هو مبين أدناه:

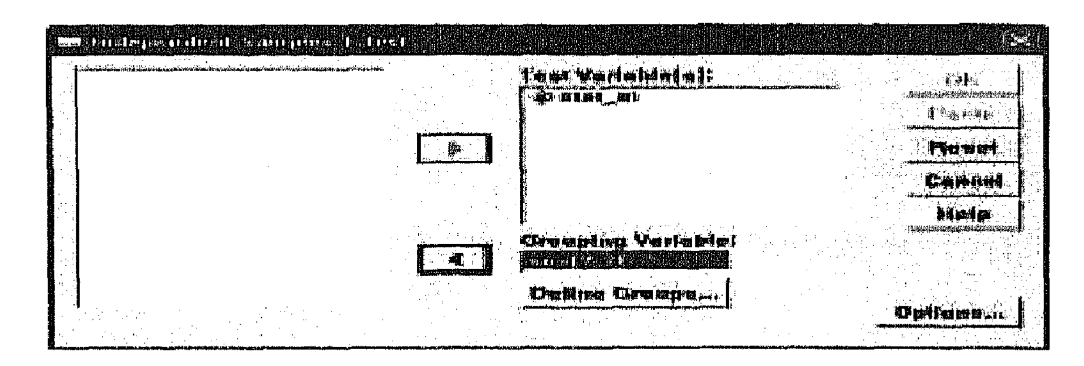
	sex	stat_m
1	1.00	16.00
2	1.00	18.00
3	1.00	15.00
4	1.00	16.00
5	1.00	12.00
6	1.00	15.00
7	1.00	18.00
8	1.00	18.00
9	1.00	18.00
10	1.00	14.00
11	2.00	16.00
12	2.00	12.00
13	2.00	17.00
14	2.00	14.00
15	2.00	19.00
16	2.00	11.00
17	2.00	10.00
18	2.00	19.00
19	2.00	12.00
20	2.00	17.00
21		

Analyze – Compare Means – Independent-Samples T Test…



ضع المتغير stat_m في خانة: Test Variable(s):

وضع المتغير sex في خانة: Grouping Variable:



انقر زر Define Groups، ضع القيمة 1 في :Group1، وضع القيمة 2 في :Group2، ثم اضغ زر Continue، ثم اضغط زر Ok

Define Groups

- Use specified values
  - Group 1: 1
  - Group 2: 2
- Cut point:

Continue | Cancel | Help

تظهر لدينا المخرجات أدناه:

الإحصاءات الوصفية لكل عينة

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
STAT_M	1.00	10	16.0000	2.0548	.6498
	2.00	10	14.7000	3.3350	1.0546

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
STAT_M	Equal variances assumed	5.356	.033	1.049	18	.308	1.3000	1.2387	-1.3025	3.9025
	Equal variances not assumed			1.049	14.972	.311	1.3000	1.2387	-1.3407	3.9407

النتيجة:

تم اختبار تجانس التباين للفئتين بواسطة اختبار Levence Test ، حيث حسبت قيمة F وكانت = 5.356 ومستوى دلالتها فكانت = 0.033 وذلك لتحديد أي الاختبارين سنستخدم، هل سنستخدم اختبار T في حالة تساوي التباين Equal variances assumed أم اختبار T في حالة عدم تساوي التباين Equal variances not assumed

استخدم اختبار T لفحص السؤال "هل يختلف تحصيل الطلاب عن الطالبات"؟ أو "هل يرتبط تحصيل الطلاب بالجنس"؟

وقـــد وجـــد مـــن خـــلال نتائج هذا الاختبار أنه ليس هناك فرقاً بين تحصيل الطلاب والطالـــبات حيث بلغت قيمة t=1.049 وهي ليست ذات دلالة إحصائية على مستوى أقل من 0.05

نخـــتار قـــيمة t ومستوى دلالتها بناء على اختبار F لنقرر هل نختار اختبار T في حالة افتراض تساوي التباين أم اختبار T في حالة افتراض عدم تساوي التباين.

في هـــذه الحالة نختار اختبار T في حالة افتراض عدم تساوي التباين لأن مستوى دلالة α=0.033 وقيمة F=5.356 أقل من 0.05 وبالتالي فان تباين الفئتين غير متساوي.

وقد بلغ متوسط تحصيل الطلاب 16 بإنحراف معياري 2.0548 في حين بلغ متوسط تحصيل الطالبات 14.7 بإنحراف معياري 3.3350 حيث يتبين أن لا فرق في التحصيل بسبب الجنس.

ملاحظة: استخدام نقطة القطع Cut Point، والرسومات البيانية.

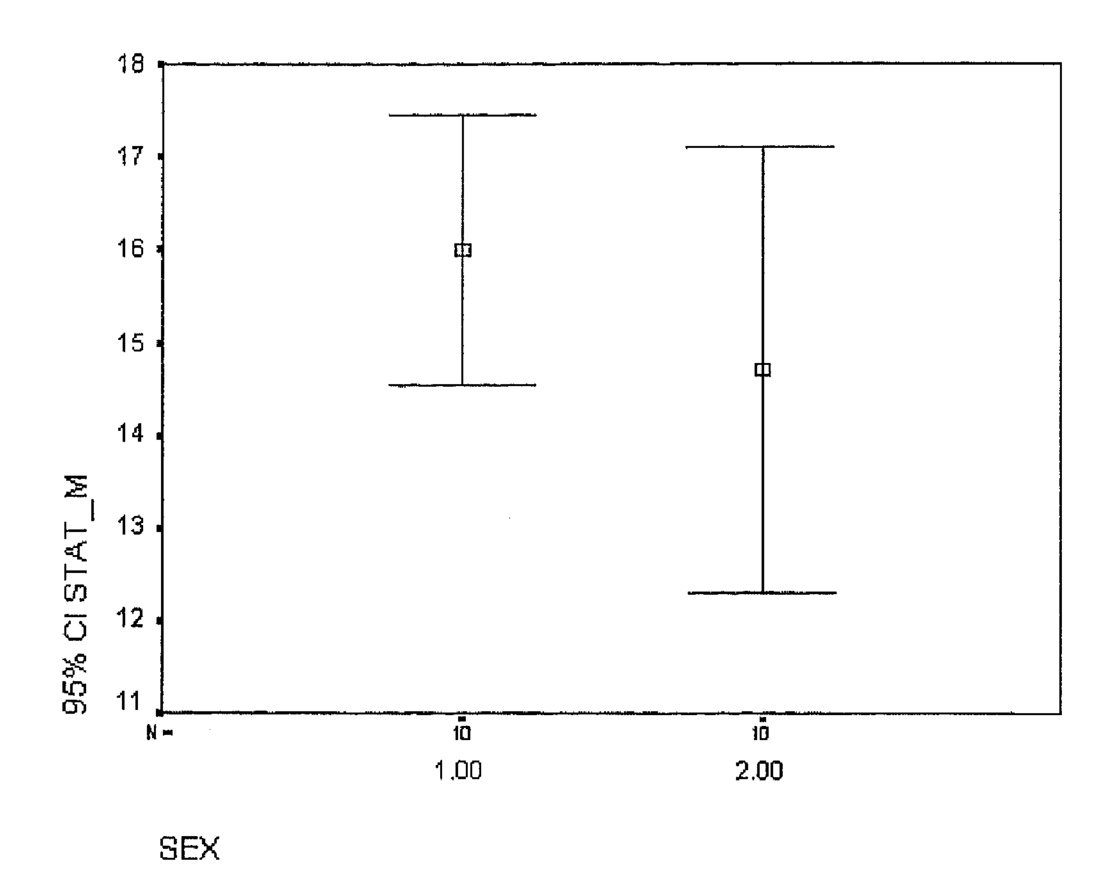
T-Test

Group Statistics

	AGE	N	Mean	Std. Deviation	Std. Error Mean
STAT_M	>= 20.00	8	14.5000	3.11677	1.10195
	< 20.00	12	15.9167	2.50303	.72256

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
STAT_M Equal variances assumed	.830	.374	-1.125	18	.275	-1.4167	1.25884	-4.06140	1.22806
Equal variances not assumed			-1.075	12.807	.302	-1.4167	1.31772	-4.26779	1.43446



4-5 تـمــاريـــن Exercise.

ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- علاقة Z ب $\sigma_{x'}$ كعلاقة T بـ ..... .

أ- σ ب- σ^2 ج- S^2x د- Sx'

2- أي من التالي يمكن اعتباره فرضية إحصائية؟

أ- $\sigma = 10$ ب- $X = 27$ ج- $r = 0.2$ د- $Sx = 8$

3- أي من التالي يمكن اعتباره فرضية إحصائية؟

أ- $\mu = 70$ ب- $X = 27$ ج- $r = 0.2$ د- $Sx = 8$

4- كم يجب أن يكون حجم العينة حتى يكون الخطأ المعياري للوسط الحسابي x' $\sigma = 10\%$ من الانحراف المعياري σ في المجتمع؟

أ- $\sqrt{10}$ ب- 1000 ج- 10 د- 100

5- إذا كانت Z=2.1 وكانت الفرضية البديلة غير متجهة، فإنه بالامكان :

أ- رفض الفرضية الصفرية على مستوى $\alpha = 0.01$

ب- رفض الفرضية الصفرية على مستوى $\alpha = 0.05$

ج- رفض الفرضية الصفرية على مستوى $\alpha = 0.05$

د- رفض الفرضية الصفرية على المستويين $\alpha = 0.01$، $\alpha = 0.05$

6- إذا كانت فترة الثقة 95% للوسط الحسابي تمتد ما بين 53.6 – 76.4 فأي من الفرضيات الإحصائية التالية تعتبر مرفوضة على مستوى $\alpha = 0.05$، مع الأخذ بالحسبان أن فرضيتها البديلة غير متجهة:

أ- $\mu = 51.6$ ب- $\mu = 65$ ج- $\mu = 75.4$ د- $\mu = 76.1$

7- أفترض أن الفرضية الصفرية التالية $H0: \mu = 60$ كانت مرفوضة على مستوى $\alpha = 0.01$ فهل هذا يعني أن القيمة 60 تقع:

أ- ضمن فترة الثقة95% ب- ضمن فترة الثقة99%

ج- خارج فترة الثقة95% د- خارج فترة الثقة99% ، 95%

8- يستخدم اختبار T كاختبار للفرضية الصفرية حول الوسط الحسابي إذا كانت .... غير معلومة.

أ- n ب- x' ج- σ د- α

9- أيهما أكثر شيوعاً في الاستخدام لفحص الفرضية حول الوسط الحسابي؟

أ- اختبار Z ب- اختبار T ج- اختبار F د- اختبار χ^2

10- تكون قيم T الحرجة عند نفس مستوى الدلالة:

أ- اكبرللاختبار بذيل. ب- اكبرللاختبار بذيلين.

ج- نفس القيمة. د- لاشيء مما ذكر.

11- ما هي قيم T الحرجة في ذيل واحد وفي ذيلين إذا كانت $n=21$ ، $\alpha = 0.05$

أ- (1.721, 2.080) ب- (2.528, 2.845)

ج- (1.725, 2.086) د- (2.518, 20831)

12- ما هي قيم T الحرجة في ذيل واحد وفي ذيلين إذا كانت $n=30$ ، $\alpha = 0.01$

أ- (2.457, 2.760) ب- (1.699, 2.045)

ج- (2.462, 2.756) د- (1.697, 2.042)

13- درست إحدى شعب الإحصاء المتقدم التي عدد أفرادها 45 طالباً وحدة اختبار الفرضيات حول الوسط الحسابي باستخدام مختبر الحاسوب، وعند إجراء اختبار يقيس مهارتهم في هذه الوحدة وجد أن متوسط أدائهم = 65.3 فإذا كانت النتائج المعلنة لجميع الطلاب الآخرين هي 50 للوسط الحسابي و 12 للانحراف المعياري، فهل هناك ما يدعو إلى القول بأن نتائج الطلبة الذين استخدموا المختبر أعلى من نتائج الآخرين؟ استخدم $\alpha = 0.01$

أ- نعم لأن $Z = 8.56$ وهي تقع في منطقة رفض الفرضية الصفرية.

ب- نعم لأن $Z = 8.56$ وهي تقع في منطقة قبول الفرضية الصفرية.

ج- لا لأن $Z = 8.56$ وهي تقع في منطقة رفض الفرضية البديلة.

د- لا لأن $Z = 8.56$ وهي تقع في منطقة قبول الفرضية البديلة.

14- فحص فادي فرضيته الصفرية $H0: \mu = 10$ بالاستعانة باختبار بذيلين $\alpha = 0.05$ مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بذيلين $\alpha = 0.05$ مستخدماً عينة حجمها 100 فإن احتمال ارتكاب الخطأ من النوع الأول يكون أكبر ل:

أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

15- فحـــص فـــادي فرضـــيته الصفرية $H0: \mu = 10$ بالاستعانة باختبار بذيلين $\alpha = 0.05$ مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بذيلين $\alpha = 0.05$ مـــستخدماً عينة حجمها 100 ، إذا كانت القيمة الحقيقية $\mu = 12$ فأيهما يكون احتمال ارتكابه للخطأ من النوع الثاني اكبر؟

أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

16- فحـــص فـــادي فرضـــيته الصفرية $H0: \mu = 10$ بالاستعانة باختبار بذيلين $\alpha = 0.05$ مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بذيلين $\alpha = 0.05$ مـــستخدماً عينة حجمها 100 ، إذا كانت القيمة الحقيقية $\mu = 12$ فأيهما يكون اختباره أقوى؟

أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

17- يدعي عميد الكلية بان توجيه إنذارات إلى الطلاب يقلل من غياب الطلاب في الجامعة. فـــاذا علم أن متوسط غياب الطلاب في الكلية في الأعوام المنصرمة كان 4.6 أيام للطالب الواحد. وبعد أن استخدم هذا الأسلوب لهذا العام مع صف عدد طلابه 25 طالباً وجد أن متوســـط الغياب كان 3.7 أيام للطالب الواحد بانحراف معياري غير متحيز = 1.2 يوم. هل تدعم هذه البيانات صحة ادعاء العميد؟ استخدم $\alpha = 0.05$

أ- نعم، لأن $T=-3.75$ وهي تقع في منطقة رفض الفرضية الصفرية.

ب- لا، لأن $T=-3.75$ وهي تقع في منطقة رفض الفرضية الصفرية.

ج- نعم، لأن $T=-3.75$ وهي تقع في منطقة رفض الفرضية البديلة.

د- لا، لأن $T=-3.75$ وهي تقع في منطقة رفض الفرضية البديلة.

18- فيما يلي علامات عشرة طلاب على مقياس الاتجاهات، جرى تطبيقه قبل وبعد مشاهدة فيلم عن العنف المدرسي:

اسم الطالب	قبل	بعد
فادي	10	14
جمال	17	31
نبيل	12	17
فايز	19	17
جمعه	10	13
محمد	11	24
احمد	20	25
صلاح	13	14
سمير	17	16
يوسف	20	15

هل تحسنت اتجاهاتهم نحو القضية بعد مشاهدة الفيلم؟ $\alpha = 0.05$

أ- لا وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $0.05 = \alpha$ للاختبار بذيل واحد.

ب- نعم وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $0.05 = \alpha$ للاختبار بذيل واحد.

ج- لا وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $0.05 = \alpha$ للاختبار بذيلين.

د- نعم وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $0.05 = \alpha$ للاختبار بذيلين.

19- درست مجموعتان مادة الإحصاء المتقدم بطريقتين مختلفتين (نظري، عملي) وعند انتهاء فترة التدريس أجري اختبار تحصيلي للمجموعتين والجدول أدناه يبين النتائج:

الانحراف المعياري غير المتحيز Sx	الوسط الحسابي 'X	عدد الأفراد n	
8	65	30	طريقة النظري
9	55	25	طريقة عملي

هل تختلف الطريقتان في تأثيرهما على التحصيل؟ علماً بأن المجموعتين قد اختيرتا بطريقة عشوائية.

أ- نعم وذلك لأن T=4.361 وهي ليست ذات دلالة إحصائية على مستوى $p<0.001$

ب- لا وذلك لأن T=4.361 وهي ذات دلالة إحصائية على مستوى $p<0.001$.

ج- نعم وذلك لأن T=4.361 وهي ذات دلالة إحصائية على مستوى $p<0.001$.

د- لا وذلك لأن T=4.361 وهي ليست ذات دلالة إحصائية على مستوى $p<0.001$

1- سحبت عينتين عشوائيتين من مجتمعين طبيعيين حجم الاولى 5 والثانية 4 والبيانات موضحة في الجدول التالي اختبر أن كانت توجد فروق جوهرية بين متوسط مجتمعيهما بمستوى دلالة $\alpha=0.05$.

العينة الاولى	العينة الثانية
3.1	2.3
4.4	1.4
1.2	3.7
1.7	8.9
3.4	

2- احدى شركات المنتجات النفطية أنتجت نوع معين مطور من مادة الكاسولين المحسن لزيادة عدد الكيلو مترات من المسافات المقطوعة. ولاختبار هذا ،سحبت عينة عشوائية من 10 سيارات وسارت باستخدام الكاسولين المحسن والعادي والجدول التالي يبين المسافات المقطوعة من قبل العشر سيارات.اختبر على مستوى 0.05 انه لا توجد فروق جوهرية وإن كانت هناك فروق قدرها بحدود الثقة السابقة في السؤال.

كاسولين	
عادي	محسن
24.9	25.7
18.8	20.5
27.7	28.4
13.0	3.7
17.8	18.8
11.3	12.5
27.6	28.4
8.2	8.1
23.1	23.1
9.9	10.4

3- الجـــدول الـــتالي يمثل بيانات عن الاجور التي تتقاضاها عينتين عشوائيتين و المسحوبة من مجتمعين طبيعين أحدهم المشتغلين لدى الدولة والثانية لدى القطاع الخاص.

الاجور لدى القطاع الخاص	الاجور لدى قطاع الدولة	
35	30	حجم العينة
35558.97	3333520	الوسط الحسابي
14940.88	15129.09	النحراف المعياري

احسب فترة ثقة 95%للفرق بين متوسطي مجتمعهما.

4- اجري اختبار في احدى المساقات ومع شعبتين مختلفتين الاولى مكونة من 40 طالباً والثانية 30 طالـــباً وجد أن متوسط علامات الطلبة في الشعبة الاولى 65 درجة بانحراف معياري قدره 10 درجة ومتوسط علامات الطلبة في الشعبة الثانية 57 درجة بانحراف معياري 6 درجـــة اختـــبر إن كان هناك فروق جوهرية بين مستوى الطلبة في الشعبتين على مستوى معنوية α=0.05 .

5- إذا كـــان متوسط الزيادة في وزن 12 فأرة بعد تغذيتها بطريقة معينة لمدة معينة هو 145 غـــم وبانحراف قياس للوسط الحسابي مقداره 2.3 غم وبمستوى احتمال 5% هل يمكن القـــول أن متوسط الزيادة في الوزن نتيجة التغذية على الطريقة لا تقل عن 150 غم اختبر ذلك.

6- شركة توزيع المحروقات ارادت تقدير الوسط الحسابي للمجتمع بحدود ثقة 95% سحبت عيـــنة عشوائية مكونة من 100 عائلة مستهلكة للسولار كان معدل استهلاكها ما يعادل 1103 غالون بانحراف معياري 327.8. احسب الوسط الحسابي للمجتمع الذي سحبت مـــنه العيـــنة ثم فسر إذا استهلكت عائلة منفردة 800 غالون هل يمكن اعتبار ذلك ممكن اختبر ذلك.

7- احـــدى شركات الاتصالات اجرت بحثا حول المكالمات الطويلة فوجد ان معدل ما يدفعه المـــواطن للمكالمـــة الطويلة يساوي 17.10 دولار في الشهر. وبانحراف معياري 9.80 دولار. سحبت عينة عشوائية لـ50 قائمة تلفون.

- أوجد احتمال أن تكون مدة المكالمة اكبر من 20 دولار.
- اختبر أن $H0: \mu = 21$ ضد $H1: \mu > 21$ مستخدماً $\alpha = 0.05$

8- أجـــريت مقارنة للأسعار في مدينتين مثل اليابان وامريكا فإذا كان سعر المفرق لبعض المواد التجارية في كل من الدولتين موضحة في الجدول أدناه:

اليابان	أمريكا	
n2=30	N1=50	حجم العينة
1224.3	1154.5	معدل سعر
1843	1989	الانحراف المعياري

اختبر أن فروقات الاسعار هي أكبرمن 200 دولار بمستوى دلالة $\alpha = 0.05$

9- احسب التقدير النفطي لعينة عشوائية حجمها 12 مسحوبة من مجتمع μ

	المجموعة الأولي	المجموعة الثانية
الوسط الحسابي	38.75	35015
الانحراف المعياري	3.2	2.7
حجم العينة	100	100

وكانت أفراد العينة هي:

X: 9 6 5 3 4 7 8 9 10 3 12 6

10- في إحدى المدارس الأساسية سحبت عيينة عشوائية من الطلبة الذين سيعملون النظارات الطبـــية حجمها 20 طالباً فوجد أن 6 منهم يستخدم النظارات الطبية. فما تقديرك للسته الذين يستعملون النظارات الطبية في تلك المدرسة.

11- سحبت عينة عشوائية حجمها 400 مفردة من مجتمع انحرافة القياسي 30 ومعدل 160 احسب فترة ثقة 95% للوسط الحسابي للمجتمع الذي سحبت منه هذه العينة.

12- سحبت عينة عشوائية من إحدى المصانع الكهربائية حجمها 25 مصباحا فكان الوسط الحسابي لأعمار هذه المصابيح 890 ساعة. احسب فترة ثقة 90% لمعدل أعمار المصابيح المنتجة في هذا المصنع على أن الانحراف المعياري لإنتاجية هذا المصنع هو35 ساعة.

13- سحبت عينة عشوائية في إحدى مصانع الخيوط حجمها 60 خيطا فوجد أن معدل قوة هــذه الخيوط 90.4 كغم بانحراف معياري 7.5 كغم. أوجد فترة ثقة 98% لمعدل قوة جميع الخيوط التي ينتجها ذلك المصنع.

14- ســحبت عيــنة عشوائية مكونة من 25 طالبا من جامعة مؤته لتقدير معدل المصروف الشهري لهم. فوجد أن مصروفهم الشهري بالدينار الأردني كما يلي:

38 51 49 38 36 35 44 50
43 41 44 38 33 45 50 51
30 50 45 40 30 44 39 49 52

- ما هو تقديريك لمعدل المصروف الشهري لجميع طلبة جامعة مؤته.

- احسب فترة ثقة 95% لمعدل المصروف إذا كان المصروف يخضع لتوزيع طبيعي.

15- سجلت قياسات الحموضة (PH) لعينات من ماء المطر في 10 مواقع في منطقة صناعية فكانت 3.6 4.1 3.9 4.8 3.2 4.5 3.8 5.1 3.1 3.9

احسب فترة ثقة 95% لمعدل حموضة ماء المطر لكل المناطق.

16- إذا كانت أجور مندوبي المبيعات لكل من الذكور والإناث تخضع لتوزيع طبيعي متباين 100 للذكور و144 للإناث علما بأن المجتمعين للذكور والإناث مستقلين عن بعضهما. سحبت عينة عشوائية حجمها 170 وانحراف معياري 9 وعينة عشوائية ثانية (إناث) حجمها 24 بوسط 162 وانحراف معياري 10.

- أوجد فترة ثقة 98% لكل من μ_1, μ_2
- احسب فترة ثقة للفرق.

17- الجدول التالي يبين التحصيل العلمي لمجموعة من الطلبة في مدينتين مختلفتين.

المدينة	حجم العينة	الوسط الحسابي	الانحراف المعياري
أ	200	79	11
ب	150	73	12

احسب فترة ثقة 95% للفرق بين معدلي تحصيل الطلبة.

18- الجدول التالي يمثل الأجور التي يتقاضاها عينتين عشوائيتين والمسحوبة من مجتمعين إحداهما المشتغلين لدى القطاع الخاص والتالية لدى الدولة.

العينة الأولى(الخاص)	الثانية(الدولة)	
35	30	حجم العينة
35558.79	33335.20	معدل الأجر بالدولار
14940.88	15129.09	الانحراف المعياري

احسب فترة ثقة 95% للفرق بين متوسطي مجتمعهما.

19- البيانات التالية تمثل فروقات قياس ضغط الدم قبل وبعد تناول الدواء، فروقات ضغط الدم تساوي = 13 -2 -1 6 4 احسب الوسط الحسابي لمجتمع الفروقات بمعدل ثقة 90%.

20- تم اختيار مجموعتين من الطلبة في مادة الرياضيات وسحبت عينتين عشوائيتين من طلاب وطالبات إحدى المدارس وكانت النتائج كما يلي:

	طلاب	طالبات
الوسط الحسابي	18	85
الانحراف المعياري	5	4
حجم العينة	10	12

علـــى افتـــراض أن المجتمعين يتوزعان قريبا من التوزيع الطبيعي وأن تباينهما غير معلوم ولكنهما متساويان احسب فترة ثقة 97% للفرق بين متوسط المجتمعين.

الفصل السادس

اختبار الفرضيات حول التباينات

Testing Hypothesis Inference Regarding Variances

1-6 اختبار فرضية تتعلق بالتباين لمجتمع واحد.
2-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين مستقلين.
3-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين غير مستقلين.
4-6 تمارين Exercise.

الفصل السادس

اختبار الفرضيات حول التباينات

Testing Hypothesis Inference Regarding Variances

1-6 اختبار فرضية تتعلق بالتباين لمجتمع واحد.

Testing Hypothesis Inference Regarding Variances

اختبار الفرضيات حول التباين ضروري في الاختبارات الإحصائية التي تتطلب توفر تجانس التباين.

- توزيع كاي تربيع (كا^2) χ^2

- لا يأخذ قيماً سالبة.
- القيمة العظمى للاحتمال فيه(68%من الحالات) تنحصر بين 0 – 1.
- ملتو التواء موجب، وإن الالتواء يقل بإزدياد درجات الحرية.
- ليس له شكل محدد ولكن يعتمد شكله على درجات الحرية.

$$\chi^2 = \sum \frac{(Xr - \mu)^2}{\sigma^2}$$

X : المشاهدة على المتغير التابع.

μ : وسط المجتمع الإحصائي على المتغير التابع.

σ^2 : التباين في المشاهدات للمجتمع على المتغير التابع.

مـــثال6–1) استخرج قيمة χ^2 التي تحصر 0.05 من المساحة في الذيل الموجب عندما تكون درجات الحرية = 19

الحـــل: مـــن جدول χ^2 نبحث عن تقاطع 0.05 من المساحة مع درجات الحرية 19 فنجد قيمة $\chi^2 = 30.14$

استخدام توزيع χ^2 في اختبار الفرضيات حول التباين

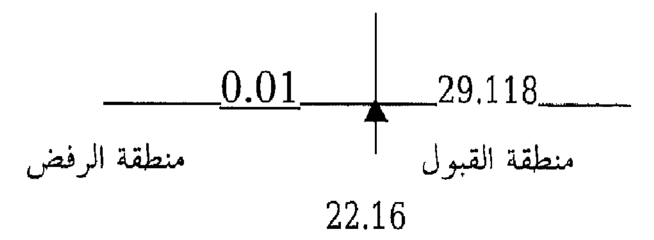
$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2}$$

S: التقدير غير المتحيز للانحراف المعياري   n : عدد المشاهدات.

σ: الانحراف المعياري للمجتمع الاحصائي.   γ=n-1

مـــثال6-2) إذا كان الانحراف المعياري في أطوال الرجال في مجتمع معين الذين تتراوح أعمـــارهم بـــين 20-25 ســـنة يساوي 2.5 إنش، ويراد معرفة فيما لو كان هذا الانحراف المعياري هو نفسه للنساء من نفس العمر، فتم اختيار 41 إمرأه ممن تتراوح أعمارهن بين 20-25 سنه وجرى قياس أطوالهن فكان الانحراف المعياري لهن 2.133 إنش.

استخدم هذه البيانات في فحص ادعاء الباحث بأن التباين في اطوال النساء يقل عما هو الحال عند الرجال من نفس العمر استخدم مستوى الدلالة $\alpha = 0.01$



الحل:

$H0: \sigma^2 = (2.5)^2 = 6.25$
$H1: \sigma^2 < 6.25$

بمـــا ان مستوى الدلالة$\alpha = 0.01$ نبحث في جدول χ^2 تحت عمود 0.99 ودرجات حرية 41-1= 40 نجد ان القيمة الحرجة = 22.16

القيمة المحسوبة

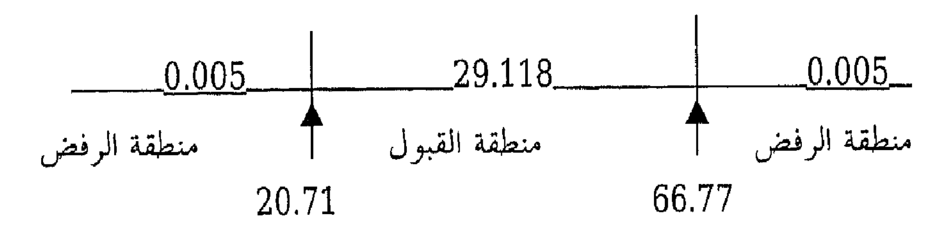
$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} = \frac{(41-1)(2.133)^2}{6.25} = 29.118$$

النتيجة:

بمـــا ان القيمة المحسوبة 29.118 أكبر من القيمة الحرجة = 22.16 لا نستطيع رفض الفرضية الصفرية، لأن منطقة الرفض للفرضية الصفرية في الجهة اليسرى.

لـــذلك لا يـــوجد ما يكفي لتدعيم فرضية الباحث بأن التباين في اطوال النساء ذوات الاعمـــار من (20 إلى 25) يقل عن التباين في اطوال الرجال من نفس العمر على مستوى الدلالة $\alpha = 0.01$

دعنا نفرض أن الفرضية البديلة كانت غير متجهة



$$H0: \sigma^2 = (2.5)^2 = 6.25$$
$$H1: \sigma^2 \neq 6.25$$

بما ان الفرضية البديلة غير متجهة تكون منطقة الرفض على الذيلين كما يلي

القيم الحرجة لمنطقة القبول $\chi^2_{40\,،0.005} = 66.77$ والقيمة $\chi^2_{40\,،0.995} = 20.71$

القيمة المحسوبة

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} = \frac{(41-1)(2.133)^2}{6.25} = 29.118$$

بمـــا ان القيمة المحسوبة 29.118 تقع ضمن منطقة قبول الفرضية الصفرية لا نستطيع رفض الفرضية الصفرية.

لـــذلك لا يـــوجد ما يكفي لتدعيم فرضية الباحث بأن التباين في اطوال النساء ذوات الاعمار 20 على 25 يختلف عن التباين في اطوال الرجال من نفس العمر على مستوى الدلالة $\alpha = 0.01$

فترة الثقة للتباين: فترة الثقة 0.95% تكتب

$$0.95 = \frac{(n-1)S^2}{\chi^2_{(n-1,0.975)}} \geq \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{(n-1,0.025)}}$$

تعني أن الاحتمال = 0.95 بأن تقع القيمة الصحيحة أو الحقيقية للتباين بين الحدين

$\frac{(n-1)S^2}{\chi^2_{(n-1,0.975)}}$ كحد أعلى ، $\frac{(n-1)S^2}{\chi^2_{(n-1,0.025)}}$ كحد أدنى.

مثال6-3: افرض ان عينة من 25 شخصاً تم اختيارهم بشكل عشوائي من مجتمع يتخذ شكل التوزيع الطبيعي في الخاصية تحت الدراسة. وعند حساب التباين غير المتحيز لتلك الخاصية في أفـــراد العيـــنة، وجـــد أنه =16، كون فترة الثقة 95% للتباين في هذه الخاصية في المجتمع الإحصائي.

الحل:

بالرجوع إلى جدول(χ^2) نجد أن $\chi^2_{(24,0.975)} = 12.40$ وأن $\chi^2_{(24,0.025)} = 39.36$

وبذلك يكون الحد الأدنى = $\frac{16*24}{39.36} = 9.76$

وبذلك يكون الحد الأعلى = $\frac{16*24}{12.40} = 30.97$

توزيع ف

يستخدم لاختبار تساوي التباين في مجتمعين احصائيين

$$F = \frac{\chi^2 \gamma_1/\gamma_2}{\chi^2 \gamma_2/\gamma_1}$$

ف مـــتغير عشوائي يتكون من النسبة لقيمتين مستقلتين من قيم χ^2 كل منهما مقسومة على درجة حريتها.

2-6 اختــبار فرضــية تــتعلق بتــساوي التــباين لمجتمعــين مستقلين.

Testing Hypothesis about Two Independent Variances

$H0: \sigma^2_1 = \sigma^2_2$

$F = \underline{S^2_1}$

S^2_2

$\gamma_1 = n_1 - 1$ درجات الحرية للبسط

$\gamma_2 = n_2 - 1$ درجات الحرية للمقام

يمكن استخدام اختبار F بذيل أو بذيلين.

الفرضـــية متجهة: إذا استخدم بذيل واحد تكون القيم الحرجة للرفض أو القبول مباشرة في جدول F.

الفرضية غير متجهة: يجب أن نقسم $\alpha/2$ ونستخرج قيمتين حرجتين من قيم F هما:

$F(\gamma_1,\gamma_2, \frac{\alpha}{2})$ كحد أدنى

$F(\gamma_1,\gamma_2,1-\frac{\alpha}{2})$ الحد الأعلى ويستخرج مباشرة من الجدول

$$\frac{1}{F(\gamma_2,\gamma_1,1-\frac{\alpha}{2})} = F(\gamma_1,\gamma_2, \frac{\alpha}{2})$$ الحد الأدنى يستخرج من العلاقة

مثال6-4) اعتقد باحث أن الطلاب الاذكياء يبدعون في التحصيل في جو تنافسي بينما يقـــل تحـــصيلهم نسبياً في جو تعاوني وبذلك يزداد التباين في التحصيل اذا تم التدريس في جو تنافسي ولكنه يقل إذا تم تدريسهم في جو تعاوني. ولأختبار هذه الفرضية قام باختبار مجموعتين بـــشكل عـــشوائي درست الأولى في جو تنافسي والثانية في جو تعاوني. وعند انتهاء التجربة استخرج الانحراف غير المتحيز لهاتين المجموعتين فكان لمجموعة الجو التنافسي S1=6.4 ولمجموعة الجو التعاوني S2=3.9 فإذا كان عدد أفراد المجموعة الاولى n1=17 وعدد أفراد المجموعة الثانية n2=15 فهل تدعم هذه البيانات صحة إدعاء الباحث؟

واذا كان ادعاؤه هو ان اختلاف الجو التدريسي لا يؤثر على التباين في التحصيل، فهل تدعم هذه البيانات صحة مثل هذا الادعاء؟ استخدم في الحالتين $\alpha=0.05$

في حالة الفرضية متجهة:

$$H0: \sigma^2_1 = \sigma^2_2$$
$$H1: \sigma^2_1 > \sigma^2_2$$

بما ان البيانات مستقلة إذا الاختبار المناسب هو F

$\gamma_1 = n_1 - 1$ درجات الحرية للبسط $= 17 - 1 = 16$

$\gamma_2 = n_2 - 1$ درجات الحرية للمقام $= 15 - 1 = 14$

نستخرج قيمة المختبر الإحصائي $F = \frac{S^2_1}{S^2_2} = \frac{(6.4)^2}{(3.9)^2} = 2.69$

القيمة الحرجة F هي $F_{(16,14,0.95)} = 2.44$

بما أن القيمة المحسوبة (F=2.69) اكبر من القيمة الحرجة ($F_{(16,14,0.95)} = 2.44$)
لذلك نرفض الفرضية الصفرية ونستنتج أن تدريس الطلاب في جو تنافسي يؤدي الى جعل التباين في علاماتهم أعلى وبدلالة احصائية غير مستوى $\alpha=0.05$ مما هي عليه الحال في الجو التعاوني.

في حالة الفرضية غير متجهة:

$$H0: \sigma^2_1 = \sigma^2_2 \qquad H1: \sigma^2_1 \neq \sigma^2_2$$

القيمة المحسوبة (F=2.69)

نستخرج قيمة المختبر الإحصائي $F = \frac{S^2_1}{S^2_2} = \frac{(6.4)^2}{(3.9)^2} = 2.69$

نستخرج قيمة المختبر الإحصائي $F = \frac{S^2_2}{S^2_1} = \frac{(3.9)^2}{(6.4)^2} = 0.37$

القيمتين الحرجتين للمختبر الاحصائي F وهما:

الحد الأعلى يستخرج مباشرة من الجدول $F(\gamma_1,\gamma_2,1-\frac{\alpha}{2}) = F_{(16,14,0.975)} = 2.95$

الحد الأدنى يستخرج من العلاقة

$$\frac{1}{F(\gamma_2,\gamma_1,1-\frac{\alpha}{2})} = F(\gamma_1,\gamma_2,\frac{\alpha}{2}) = F_{(16,14,0.025)} = 1/F_{(14,16,0.975)} = 1/2.75 = 0.36$$

بما أن القيمة المحسوبة ل F هي: 0.37 و 2.6 تقع ضمن منطقة القبول بين الحدين 2.95 و 0.36 ولذلك نقبل الفرضية الصفرية ونستنتج أن تدريس الطلاب في جو تنافسي لا يؤدي الى جعل الاختلاف في التباين في علاماتهم عن التباين في علامات الذين يدرسون في جو تعاوني ذا دلالة احصائية.

3-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين غير مستقلين.

Testing Hypothesis about Two Dependent Variances

$H0: \sigma^2_1 = \sigma^2_2$

$\gamma = n - 2$ درجات الحرية

$$T = \frac{S_1^2 - S_2^2}{\sqrt{\frac{4S_1^2 - S_2^2}{n-2} * (1 - r^2)}}$$

S^2_1 : التباين غير المتحيز للعينة الأولى.

S^2_2 : التباين غير المتحيز للعينة الثانية.

n : عدد أزواج المشاهدات.

r : معامل الارتباط بين المشاهدات في العينة الأولى والمشاهدات في العينة الثانية.

مثال6-5: اعتقد باحث أن التجانس في الأداء على مقياس للاتجاهات نحو المدرسة يزداد مع تقدم الطالب في المستوى التعليمي. وبمعنى آخر يقل التباين كلما ارتفع الطالب من مستوى تعليمي لآخر. وللتحقق من صحة هذا الادعاء أجرى اختبارا للاتجاهات على عينة من الطلاب من مستوى الثالث الاعدادي ثم أعاد إجراء الاختبار على نفس العينة وهم في مستوى الثالث الثانوي. وقد كان عدد الذين اخذوا الاختبار في المرتين يساوي 74 طالباً. وكان التباين في العلامات في الاجراء الأول $S^2_1=105$ وفي الاجراء الثاني أصبح التباين $S^2_2=82$ ، أما معامل الارتباط بين علامات الطلاب في الاجرائين فكان $r=0.8$ هل تدعم هذه البيانات صحة إدعاء الباحث؟ استخدم $\alpha = 0.05$

الحل:

$H0: \sigma^2_1 = \sigma^2_2$
$H1: \sigma^2_1 > \sigma^2_2$

σ^2_1 : التباين في الثالث الإعدادي.

σ^2_2 : التباين في الثالث الثانوي.

القيمة المحسوبة للمختبر الاحصائي T

$$T = \frac{S^2_1 - S^2_2}{\sqrt{\frac{4 S^2_1 S^2_2}{n-2} * (1 - r^2)}}$$

$$T = \frac{105 - 82}{\sqrt{\frac{4 * 105 * 82}{74 - 2} * (1 - 0.8^2)}} = 1.752$$

القيمة الحرجة $T_{\gamma,1-\alpha}$

$\gamma = 74 - 2 = 72$ درجات الحرية

One Tail الاختبار بذيل واحد

$\alpha = 0.05$ مستوى الدلالة

$T_{\gamma,1-\alpha} = T_{72,1-0.05} = T_{72,0.95} = 1.671$

النتيجة:

بما أن القيمة المحسوبة للمختبر الاحصائيT (1.752) أكبر من القيمة الحرجة (1.671) نرفض الفرضية الصفرية على مستوى $\alpha = 0.05$ ، ونستنتج أن البيانات تدل على أن التباين في إتجاهات الطلاب نحو المدرسة يقل مع ارتفاع مستواهم التعليمي.

4-6 تمـــاريـــــن Exercise.

تمرين (1) : ضع رمز الإجابة (نعم أو لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- يمكـــن إهمـــال افتـــراض تجـــانس التباين في اختبار T اذا تساوى عدد الأفراد في كل من المجموعتين.

2- يتم اختبار الفرضية $H0: \sigma^2 = C$ باستخدام اختبار T

3- يتم اختبار الفرضية $H0: \sigma^2{}_1 = \sigma^2{}_2$ مشاهدات مستقلة باستخدام اختبار F

4- يتم اختبار الفرضية $H0: \sigma^2{}_1 = \sigma^2{}_2$ مشاهدات غير مستقلة باستخدام اختبار T

5- تصاغ الفرضيات الاحصائية بعد أن يتم جمع البيانات.

6- اختبار التباين بصورة عامة لا تتأثر بعدم الايفاء بافتراضات التوزيع الطبيعي.

7- اذا كانـــت الفرضـــية البديلة $H1: \sigma^2{}_1 > \sigma^2{}_2$ فإن القيمة الحرجة للاحصائي F هي التي يعطيها جدول F مباشرة؟

8- اذا كانت n1=n2 فلماذا نهتم باختبار الفرضية $H0: \sigma^2{}_1 = \sigma^2{}_2$ ؟

تمرين (2) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- أن توزيع χ^2 يتصف بما يلي:

أ- توقع التوزيع=درجات الحرية.

ب- يكون دائماً متماثلاً.

ج- يكون ملتوياً التواء موجباً.

د- يكون ملتوياً التواء سالباً.

2- أن توزيع T يتصف بما يلي:

أ- توقع التوزيع=درجات الحرية. ب- يكون دائماً متماثلاً.

ج- يكون ملتوياً التواء موجباً. د- يكون ملتوياً التواء سالباً.

3- أن توزيع F يتصف بما يلي:

أ- توقع التوزيع=درجات الحرية. ب- يكون دائماً متماثلاً.

ج- يكون ملتوياً التواء موجباً. د- يكون ملتوياً التواء سالباً.

4- إن القيمة الحرجة ($F_{5,7,0.05}$) هي:

أ- 3.95 ب- 7.46 ج- 4.88 د- 16.45

5- ان القيمة الحرجة ($\chi^2_{5,0.99}$) هي:

أ- 0.554 ب-0.752 ج- 1.145 د- 1.610

6- ان القيمة الحرجة ($T_{24,0.975}$) هي:

أ- 2.060 ب- 2.069 ج- 2.064 د- 1.711

7- يتم اختبار الفرضية $H_0: \sigma^2 = C$ باستخدام اختبار:

أ- Z ب- T ج- F د- χ^2

8- يتم اختبار الفرضية $H_0: \sigma^2_1 = \sigma^2_2$ مشاهدات مستقلة باستخدام اختبار........

أ- Z ب- T ج- F د- χ^2

9- عينتان مستقلتان، التباين في المشاهدات في أحد المتغيرات في الأولى =100 وعدد أفرادها= 15 والتباين في المشاهدات في نفس المتغير للعينة الثانية =64 وحجمها = 36، هل يختلف التباين في المجتمعين احصائياً؟

أ- لا، لأن F=1.56 وهي ذات دلالة احصائية.

ب- نعم، لأن F=1.56 وهي ذات دلالة احصائية.

ج- نعم، لأن F=1.56 وهي ليست ذات دلالة احصائية.

د- لا، لأن F=1.56 وهي ليست ذات دلالة احصائية.

10- يمكن إهمال افتراض تجانس التباين في اختبار T اذا :

أ- عدد الأفراد في المجموعة الأولى يساوي عدد الأفراد في المجموعة الثانية.

ب- عدد الأفراد في المجموعة الأولى أكبر من عدد الأفراد في المجموعة الثانية

ج- عدد الأفراد في المجموعة الأولى أقل من عدد الأفراد في المجموعة الثانية.

د- لاشيء مما ذكر.

11- أي من التالية ليس من خصائص توزيع كاي تربيع (كا^2) χ^2

أ- يأخذ قيماً سالبة.

ب- القيمة العظمى للاحتمال فيه(68%من الحالات) تنحصر بين 0 – 1.

ج- ملتو التواء موجب، وإن الالتواء يقل بازدياد درجات الحرية.

د-ليس له شكل محدد ولكن يعتمد شكله على درجات الحرية.

12- يتم اختبار الفرضية $H_0: \sigma^2_1 = \sigma^2_2$ مشاهدات غير مستقلة باستخدام اختبار......

أ- Z　　ب- T　　ج- F　　د- χ^2

13- في توزيع χ^2 إن القيمة العظمى للاحتمال التي تنحصر بين 0 – 1 هي:

أ- 0.34　　ب-0.68　　ج- 0.95　　د- 0.99

14- اذا كـــان الانحراف المعياري للمجتمع = 2.5 ،وكان الانحراف المعياري للعينة=2.13 وعدد العينة=31 ومستوى الدلالة=0.01 فإن قيمة χ^2 المحسوبة تساوي:

أ- 0.34　　ب-0.68　　ج- 0.95　　د- 0.99

15- اذا كـــان الانحراف المعياري للمجتمع = 2.5، وكان الانحراف المعياري للعينة=2.13 وعدد العينة=31 ومستوى الدلالة=0.01 فإن قيمة χ^2 الحرجة تساوي:

أ- 20.599　　ب- 16.306　　ج- 18.493　　د- 14.953

الفصل السابع

اختبار الفرضيات حول معاملات الارتباط

Hypotheses Testing Regarding Correlation Coefficients

7-1 مقدمة
7-2 اختبار الفرضية حول معامل ارتباط واحد.
7-3 اختبار الفرضيات حول الفرق بين معاملي ارتباط مستقلين.
7-4 اختبار الفرضيات حول معاملي ارتباط للبيانات غير المستقلة.
7-5 استخدام برمجية SPSS في حل المسائل.
7-6 تمارين Exercise .

الفصل السابع

اختبار الفرضيات حول معاملات الارتباط

Hypotheses Testing Regarding Correlation Coefficients

1-7 مقدمة

الارتباط Correlation

إن نظرية الارتباط تظهر قوة العلاقة بين متغيرين مع إمكانية تحديد نوع وقوة العلاقة بين الظواهـــر، كالعلاقـــة بـــين مستوى التعليم والأداء، والعلاقة بين معدل الثانوية العامة ومعدل الجامعة، والعلاقة بين المستوى الاقتصادي والتحصيل.

إن الهـــدف مـــن تحليل الارتباط Correlation هو معرفة وجود علاقة بين متغيرين أو مجمـــوعة من المتغيرات المستقلة Independent Variables (X_i) مع المتغير التابع Dependent Variable (Y_i) مـــن عدم وجودها، وهناك عدة مقاييس لتحديد درجة العلاقة والارتباط بين المتغيرات.

- **قياس الارتباط Measures of Correlation**

1- رسم شكل الانتشار Scatter Plot.

2- القياس الكمي للارتباط Quantitative Measure

معامل الارتباط الخطي البسيط Simple Linear Correlation

أ- معامل ارتباط بيرسون الخطي Person Linear Correlation Coefficient

معامل ارتباط بيرسون يقيس قوة واتجاه العلاقة الخطية فقط بين متغيرين كميين.

$$\text{ر} = \frac{\sum (\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})}{\sqrt{\sum (\text{س} - \bar{\text{س}})^2 \sum (\text{ص} - \bar{\text{ص}})^2}}$$

$$r_{xy} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(x-\bar{x})^2\ \Sigma(y-\bar{y})^2}}$$

$$r_{xy} = \frac{\Sigma Z_x Z_y}{n}$$

$$ر = \frac{\sum ز_س\ ز_ص}{ن}$$

$$ر = \frac{ن(\sum س_ر\ ص_ر) - (\sum س_ر)(\sum ص_ر)}{\sqrt{ن\sum س_ر^2 - (\sum س_ر)^2}\ \sqrt{ن\sum ص_ر^2 - (\sum ص_ر)^2}}$$

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2) - (\Sigma x)^2]}\ \sqrt{[n(\Sigma y^2) - (\Sigma y)^2]}}$$

$$ر = \frac{\sum (س - \bar{س})(ص - \bar{ص})}{ن\ \sigma س\ \sigma ص}$$

$$ر = \frac{(\sum س_ر\ ص_ر) - ن * \bar{س} * \bar{ص}}{\sqrt{\sum س_ر^2 - ن * \bar{س}^2}\ \sqrt{\sum ص_ر^2 - ن * \bar{ص}^2}}$$

$$r = \frac{(\Sigma xy) - n\bar{X}\bar{Y}}{[\sqrt{(\Sigma x^2) - n\bar{X}^2}]\ [\sqrt{(\Sigma y^2) - n\bar{Y}^2}]}$$

*** الشروط الواجب توافرها لاستخدام معامل ارتباط بيرسون**

1. وجود علاقة خطية بين المتغيرين.
2. يجب أن تكون العينة عشوائية وقيم أفراد العينة مستقلة عن بعضها البعض.

*** تقييم قيمة معامل الارتباط Correlation Coefficient Evaluation**

جدول (24): تقييم قيمة معامل الارتباط حسب تصنيف

(Hinkle and Others, 1979)

الفئة	التفسير
من 0.00 – أقل من 0.30	منخفض جداً
من 0.30 – أقل من 0.50	منخفض
من 0.50 – أقل من 0.70	متوسط
من 0.70 – أقل من 0.90	عال
من 0.90 – أقل من 1.00	عال جداً

* (خليل الخليلي) تصنيف (Hinkle and Others, 1979)

ب– معامل ارتباط سبيرمان للرتب Spearman Correlation Coefficient

يستخدم معامل ارتباط سبيرمان ومعامل كندال تاو لقياس قوة الارتباط بين متغيرين ترتيبيين Ordinal.

$$ر = 1 - \frac{6 \sum ف_ر^2}{ن(ن^2 - 1)}$$

$$r = 1 - \frac{6 \Sigma d^2}{n(n^2 - 1)}.$$

حيث: d تعني الفرق بين رتبة X ورتبة Y

ج– معامل ارتباط إيتا Eta η بين متغيرين كل منهما متصل والعلاقة بينهما انحنائية.

د– معامل التوافق (م ت) Contingency Coefficient (CC)

معامل ارتباط بين متغيرين كل منهما منفصل، ولكن ليس بالضرورة أن يكون أي منهما منفصلاً ثنائياً، يمكن أن يستخدم عندما تكون عدد الفئات في أحد المتغيرين أو كليهما أثنين أو أكثر، معامل ارتباط بين متغيرين أحدهما أو كلاهما ينقسم إلى أكثر من حالتين، وقد يكون أحد أو كلا المتغيرين وصفية.

هـ- بايسيريال رتبي (ردر) Rank Biserial (r_{rb})

معامـــل ارتباط بين متغيرين أحد هذين المتغيرين ثنائي منفصل والآخر متغير رتبي. مثل علاقة نجاح أو فشل الطالب في دراسته والمستوى الاقتصادي لولي الأمر.

و- تتراشورك (رت ت) Tetrachoric (r_t)

معامل ارتباط بين متغيرين كل منهما منفصل بالتحويل.

ز- معامل الاقتران Coefficient of Association (r_A) :

يـــستخدم عـــندما تكـــون بـــيانات كلا المتغيرين X, Y أو أحدهما غير قابلة للترتيب التصاعدي أو التنازلي، وكان عدد الحالات التي فيها كل من المتغيرين هي حالتين فقط.

ح- معامل ارتباط بايسيريال (رب) Biserial (r_b)

معامـــل ارتباط بين متغيرين أحد هذين المتغيرين يقع على مقياس فئوي أو مقياس نسبي والآخر متغير ثنائي منفصل ولكن بصورة غير طبيعية.

ط- معامل ارتباط بوينت بايسيريال (ردد) Point Biserial (r_{pb})

معامـــل ارتباط بين متغيرين بحيث يكون أحد المتغيرين منفصلاً ثنائياً بصورة طبيعية مثل مـــتغير الجنس والمتغير الثاني يقع على مقياس فئوي Interval أو مقياس نسبي Ratio مثل متغير الذكاء.

$$r_{pb} = [(Y'_2 - Y'_1)/ S_y] (\sqrt{pq})$$

ي- معامل ارتباط فاي (Φ) Phi

معامـــل ارتباط بين متغيرين كل منهما منفصل ثنائي بصورة طبيعية Dichotomous، ويقع هذين المتغيرين على مقياس اسمي (Nominal).

يحسب معامل ارتباط فاي (Φ) بالمعادلة التالية:

$$\Phi = \frac{P1_{xy} - P1_x \, P1_y}{\sqrt{P1_x \, P0_x \, P1_y \, P0_y}}$$

$P1_{xy}$: نسبة الأفراد أصحاب العلامة 1 على المتغيرين.

$P1_x$: نسبة الأفراد أصحاب العلامة 1 على المتغير x.

$P1_y$: نسبة الأفراد أصحاب العلامة 1 على المتغير y.

$P0_x$: نسبة الأفراد أصحاب العلامة 0 على المتغير x.

$P0_y$: نسبة الأفراد أصحاب العلامة 0 على المتغير y.

معامل التحديد Determinant Coefficient التباين المفسر أو المشترك r^2

* أراد احـــد الباحثين أن يدرس العلاقة بين الذكاء C ، والابتكار IQ فحسب معامل ارتباط بيرسون بين أداء الطلاب على اختبار للذكاء C واختبار للابتكار IQ فكان يساوي $r_{iq,c}$ = 0.7 ، $r^2 = r * r$

2-7 اختبار الفرضية حول معامل ارتباط واحد.

لمعرفة ما اذا كانت معاملات الارتباط التي نستخرجها بين متغيراتنا ذات دلالة احصائية ام لا.

ما المقصود بعبارة "ان معامل الارتباط بين متغيرين كان ذا دلالة احصائية"؟

ان هـــذه العبارة استدلال احصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$ وهي تعـــني ان معامـــل الارتـــباط في المجتمع بين المتغيرين لا يختلف عن صفر، ورفض هذه الفرضية الـــصفرية يعني ان قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً، أي أنها لم تنشأ عن خطأ المعاينة.

يستخدم توزيع المعاينة لمعامل الارتباط بيرسون r في اختبار الفرضية الصفرية $H_0: \rho = 0$ يتخذ شكل توزيع T بدرجات حرية $\gamma = n - 2$ ، حيث n هي حجم العينة.

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

مـــثال1-14: وجد في احدى الدراسات ان معامل الارتباط بين المستوى الاقتصادي الاجتماعـــي والتحصيل = 0.4 عند استخدام عينة عشوائية حجمها = 30 طالب، فهل هذا المعامل يختلف جوهرياً عن الصفر؟

$H_0: \rho = 0$

$H_1: \rho \neq 0$

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.4}{\sqrt{\frac{1-0.4^2}{30-2}}} = \frac{0.4}{\sqrt{\frac{0.84}{28}}} = \frac{0.4}{0.1732} = 2.31$$

$T_{28,0.975} = 2.048$ القيمة الحرجة

النتيجة: نرفض الفرضية الصفرية لأن القيمة الحرجة للمختبر الإحصائي 2.048 أقل من القيمة المحسوبة 2.31، ونقول أن معامل الارتباط بين المستوى الاقتصادي الاجتماعي والتحصيل 0.4 كان ذا دلالة إحصائية، أي يختلف جوهرياً عن الصفر؟ يمكن تحديد القيمة الحرجة لمعامل الارتباط الذي نرفض أو نقبل الفرضية الصفرية وهي:

$$r_c = \frac{T}{\sqrt{T^2 + \gamma}}$$

$\gamma = n - 2$

T القيمة الحرجة للرفض أو القبول حسب درجات الحرية ومستوى الدلالة

استخدام جداول r_c مباشرة دون الحاجة لحساب T

عندما $r=0.4$ يدل على ارتباط ذي دلالة احصائية على مستوى $\alpha=0.05$ والقيمة الحرجة في الجدول هي 0.374

3-7 اختبار الفرضيات حول الفرق بين معاملي ارتباط مستقلين.
Hypotheses Testing Regarding the Difference between two independent Correlation Coefficients

$H_0: \rho_1 = \rho_2$

نستخدم اختبار Z بعد الاستعانة بعلامات فشر المعيارية.

$$Z = \frac{F_1 - F_2}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}$$

حيث:

n_1 : حجم العينة الأولى.

F_1 : علامة فشر المقابلة لمعامل الارتباط في العينة الأولى.

n_2 : حجم العينة الثانية.

F_2 : علامة فشر المقابلة لمعامل الارتباط في العينة الثانية.

مـــثال14-2: اعـــتقد باحث أن قوة العلاقة بين القدرة القرائية والتحصيل في مجتمع الاناث في جامعة الاسراء أقل منها في مجتمع الذكور في نفس المستوى. وعندما استخرج معامل الارتـــباط بين هذين المتغيرين على عينة عشوائية من الاناث من هذه الفئة مؤلفة من 40 طالبة وجده مساوياً 0.7، وبالمقابل وجد معامل الارتباط بين المتغيرين مساوياً 0.8 عندما استخرجه من خلال عينة من الذكور مؤلفة من 50 طالباً. فهل تدعم هذه البيانات صحة ادعاء الباحث؟

$H_0: \rho_1 = \rho_2$

$H_1: \rho_1 > \rho_2$

حيث:

ρ_1 : معامل الارتباط بين القدرة القرائية والتحصيل في مجتمع الذكور.

ρ_2 : معامل الارتباط بين القدرة القرائية والتحصيل في مجتمع الاناث.

وحيث أن ر$_1$ = 0.8 فإن $F_1 = 1.099$

وكذلك ر$_2$ = 0.7 فإن $F_2 = 0.867$

$$Z = \frac{F_1 - F_2}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} = \frac{0.867 - 1.099}{\sqrt{\frac{1}{40-3} + \frac{1}{50-3}}} = -1.06$$

النتيجة:

القـــيمة الحـــرجة للمختبر الاحصائي Z على مستوى $\alpha=0.05$ تساوي ‎-1.64 ومن الواضح أن القيمة المحسوبة للمختبر الاحصائي واقعة في منطقة قبول الفرضية الصفرية على هذا المـــستوى وبذلك نقبل الفرضية الصفرية، ونستدل على أن البيانات المتوفرة لا تدل على أن قوة العلاقة بين القدرة القرائية والتحصيل هي أقل في مجتمع الاناث عما هي عليه في مجتمع الذكور.

4-7 اختـبار الفرضـيات حـول معاملـي ارتـباط للبـيانات غيـر المستقلة.

Hypotheses Testing Regarding two dependent Correlation Coefficients

مـــثال: صدق التنبؤ بمعدل الطالب في نهاية المرحلة الجامعية من خلال معدله في الثانوية العامة مقابل صدق التنبؤ من خلال معدله في المدرسة على نفس افراد العينة.

$H_0: \rho_{31} = \rho_{32}$
$H_1: \rho_{31} > \rho_{32}$

ρ_{31} : معامل الارتباط في المجتمع بين المتغير3 والمتغير1

ρ_{32} : معامل الارتباط في المجتمع بين المتغير3 والمتغير2

$$T = (r_{31}-r_{32}) \sqrt{\frac{(n-3)(1+r_{12})}{2(1-r^2_{31}-r^2_{32}-r^2_{12}+2r_{31}r_{32}r_{21})}}$$

$\gamma = n - 3$

حيث:

n : حجم العينة.

r_{31} : معامل الارتباط بين المتغيرين 3، 1 في العينة.

r_{32} : معامل الارتباط بين المتغيرين 3، 2 في العينة.

r_{12} : معامل الارتباط بين المتغيرين 1، 2 في العينة.

مثال14-3: اعتقد باحث أن التحصيل يرتبط بدرجة أقوى بالقدرة القرائية منه بالقدرة الكتابية في مستوى الابتدائي. ولأجل ذلك طور اختبارين أحدهما يقيس القدرة القرائية والآخر يقـــيس القـــدرة الكتابية ثم طبقهما على عينة مؤلفة من 67 طالباً. وفي نهاية الفصل الدراسي ســـجل معدل هؤلاء في المدرسة كمؤشر للتحصيل. ولدى حساب معاملات الارتباط بين هذه المتغيرات حصل على القيم التالية:

معامل الارتباط بين التحصيل والقدرة القرائية = 0.8 r_{31}

معامل الارتباط بين التحصيل والقدرة الكتابية = 0.65 r_{32}

معامل الارتباط بين القدرة القرائية والقدرة الكتابية = 0.5 r_{12}

فهل تدل البيانات على صحة ادعاء الباحث؟

المتغير المشترك 3 هو التحصيل، المتغير 1 القدرة القرائية، المتغير 2 القدرة الكتابية.

$H_0: \rho_{31} = \rho_{32}$
$H_1: \rho_{31} > \rho_{32}$

$$T = (r_{31}-r_{32}) \sqrt{\frac{(n-3)(1+r_{12})}{2(1-r^2_{31}-r^2_{32}-r^2_{12}+2r_{31}r_{32}r_{21})}}$$

$$T = (0.8-0.65) \sqrt{\frac{(67-3)(1+0.5)}{2(1-(0.8)^2-(0.65)^2-(0.5)^2+2*0.8*0.65*0.5)}} = 2.28$$

$\gamma = n - 3 = 67-3 = 64$

$T_{64,0.95} = 1.67$

نرفض الفرضية الصفرية على مستوى $P<0.05$ ونستنتج أن البيانات لا تدل على عكس ما يدعيه الباحث. اذا كان معامل الارتباط بين التحصيل والقدرة القرائية أعلى منه بين التحصيل والقدرة الكتابية.

الأساليب الإحصائية المناسبة لدراسة العلاقة وفقا لعدد المتغيرات.

أساليب دراسة العلاقة بين متغيرين	مجموعة من المتغيرات المستقلة ومتغير تابع واحد	العلاقة بين مجموعة من المتغيرات المستقلة ومجموعة من المتغيرات التابعة	العلاقة بين متغيرين مع ضبط الثالث
Person's Correlation Rank- Correlation Sperman Rho Kendall's tau Biserial Correlation Widespread Biserial Correlation Point-Biserial Correlation Tetrachoric Correlation Phi Coefficient Contingency Coefficient Correlation ratio	Multiple Linear regression Discriminate Function	Canonical Correlation	Partial Correlation Part correlation

أساليب حساب العلاقة المناسبة وفقا لمستوى القياس للمتغيرين:

المتغير الأول	المتغير الثاني	المقاييس المناسبة
فتري أو نسبي	فئوي أو نسبي	معامل بيرسون (Pearson product Moment Correlation . r.) Coefficient –حاصل ضرب العزوم- (إذا كانت العلاقة خطية). نسبة الارتباط Correlation Ratio (معامل ايتا) إذا كانت العلاقة غير خطية)

رتبي	رتبي	معامل سبيرمان لارتباط الرتب Spearman's Rank Correlation Coeficient إذا كان المطلوب قياس الاقتران ووزن الرتب بميزان فتري. معامل كاندل تو لارتباط الرتب Kedall's Tau Rank Coefficient (لقياس الاقتران مع عدم وزن الرتب بميزان فتري). معامل الاقتران لجودمان وكروسكال Gooddman and Kruskal's Coefficient of Ordinal Association.
اسمي	اسمي	معامل التنبؤ المتماثل لجتمان معامل فاي ، معامل الاقتران ليول، معامل التجمع ليول (عندما يشمل كل متغير على قسمين). معامل الاقتران لبيرسون. معامل التوافق –التصاحب- Contingency عندما يكون أحد المتغيرين أو كليهما متعدد الفئات. معامل تراشورك Tetrachoric Coefficient : يستخدم إذا كان مستوى القياس في المتغيرين متصلة ثم حولت إلى اسمية.
اسمي	رتبي	معامل وليكوكسون للاقتران (إذا لم يكن هناك تمييز بين المتغير المستقل والتابع) . معامل وليكوكسون لإشارات الرتب إذا كان هناك تمييز بين المتغير المستقل والتابع. (هناك معامل خاص عندما يتكون المتغير الاسمي من قسمين، و أخر عندما يتكون المتغير الاسمي من اكثر من قسمين). رتب بايسيريال Rank Biserial.
اسمي	فئوي أو نسبي	نسبة الارتباط (مع افتراض التوزيع الاعتدالي للبيانات، وان يكون المتغير التابع هو المتغير الفتري. بوينت بايسيريال Point Biserial Correlation . بايسيربال Biserial Correlation (عندما يكون المتغير الاسمي أصلا متصلا ولكنه حول إلى اسمي كتحويل درجة مفهوم ذات سالب و موجب والتعامل معها كمتغير اسمي أو ثنائي..
رتبي	فئوي أو نسبي	معامل الارتباط المتسلسل المتعدد لجاسبن Jaspen Coefficient of Multi-serial Correlation (شرط اعتبار المتغير الرتبي متغير متصل يأخذ التوزيع الاعتدالي). معامل الارتباط الثنائي المتسلسل Biserial Correlation

ثنائي (0-1) Dichotomous	فئوي أو نسبي	معامل فاي . معامل الارتباط الثنائي المتسلسل(عندما يكون المتغير الثنائي غير حقيقي ويكون المطلوب تقدير معامل الارتباط كما لو كان المتغير متصلا). معامل ارتباط بيرسون (عندما يكون المتغير الثنائي متغيرا حقيقيا).
ثنائي	ثنائي	معامل الارتباط الثنائي المتسلسل Point Biserial Coefficient Correlation (عندما يكون المتغير الثنائي غير حقيقي ويكون المطلوب تقدير معامل الارتباط كما لو كان المتغير متصلا). معامل الارتباط الرباعي الحقيقي (فاي) Fourfold Phi Correlation (الثنائية غير حقيقية واعتبارها متصلا). معامل الارتباط الرباعي Tetrachonic Correlation. معامل ارتباط بيرسون (عندما يكون المتغير الثنائي متغيرا حقيقيا).
فئوي	فئوي	الانحدار الخطي (عند التميز بيم المتغير المستقل والتابع، العلاقة خطية، الهدف التنبؤ).
		الانحدار المنحني (عند التمييز بين المتغير التابع والمستقل، العلاقة غير خطية، الهدف التنبؤ).
		نسبة الارتباط (عندما لا يكون هناك تمييز بين المتغير المستقل والتابع، علاقة غير خطية، ليس الاقتران هدفا للقياس.

5-7 استخدام برمجية SPSS في حل المسائل.

مثال9: إذا كان غرض الباحث هو حساب معامل الارتباط بين جنس المفحوص (sex) ونـــوع الإجابـــة عن سؤال من بديلين (Agreement) لمجموعة من 10 طلاب من الجنسين، وكانت الإجابات كما يلي:

	st_no	sex	agrement
1	1	1	1
2	2	1	1
3	3	1	0
4	4	0	1
5	5	0	0
6	6	1	0
7	7	1	1
8	8	1	1
9	9	0	0
10	10	0	0

Analyze – Descriptive Statistics – Crosstabs…

ضع متغير sex في نافذة Row(s): وضع متغير agreement في نافذة Column(s): ثم اضغط زر Statistics تظهر الشاشة أدناه:

نختار Phi and Cramer's V ثم نضغط زر Continue ثم زر Ok فتظهر شاشة المخرجات أدناه:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SEX * AGREMENT	10	100.0%	0	.0%	10	100.0%

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.408	.197
	Cramer's V	.408	.197
N of Valid Cases		10	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

التعليق: يشير معامل الارتباط إلى وجود علاقة ايجابية بين جنس الطالب ونوع الإجابة، حيث تفسر أن الطلاب الذكور يميلون إلى الموافقة نحو فكرة معينة بينما تميل الإناث إلى عدم الموافقة.

مثال9-1) إذا علمت أن علامة طالب في مادة الإحصاء Stat هي X وعلامته في مادة الرياضيات Math هي y . فهل تعتقد بوجود علاقة بين العلامتين؟

X = 5 , 4 , 2 , 1
Y = 7 , 6 , 3 , 4

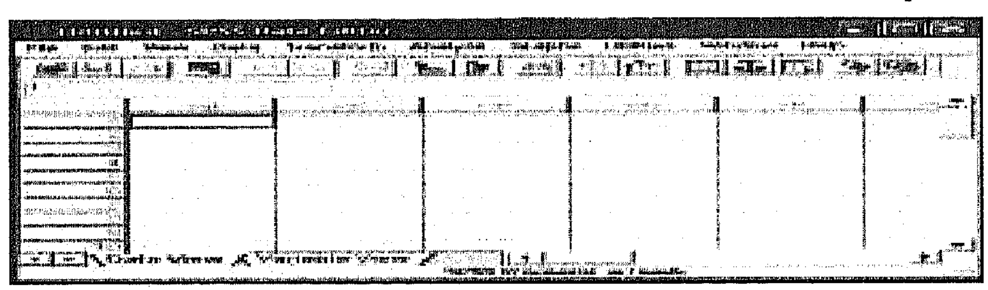
1- هل تعتقد بوجود علاقة بين العلامتين؟ ولماذا؟

2- احسب معامل ارتباط بيرسون بين هذين المتغيرين؟

3- احسب معامل ارتباط سبيرمان للرتب بين هذين المتغيرين؟

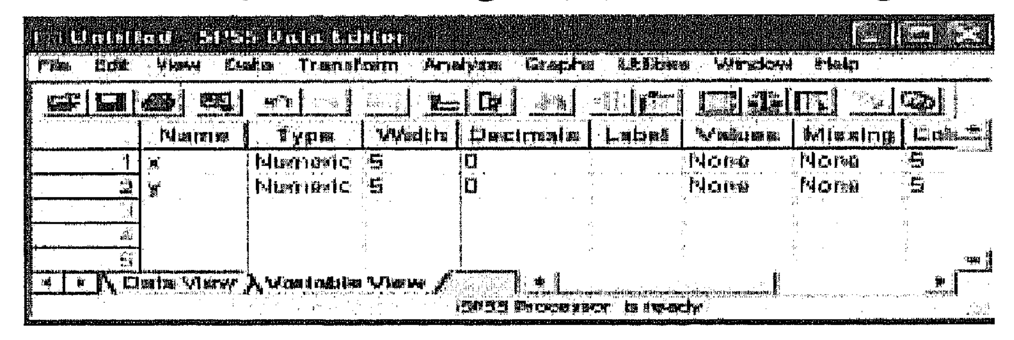
Start– Programs–SPSS for Windows–SPSS10.0 for Windows–Type in data– Ok

تظهر لديك الشاشة المبينة أدناه:

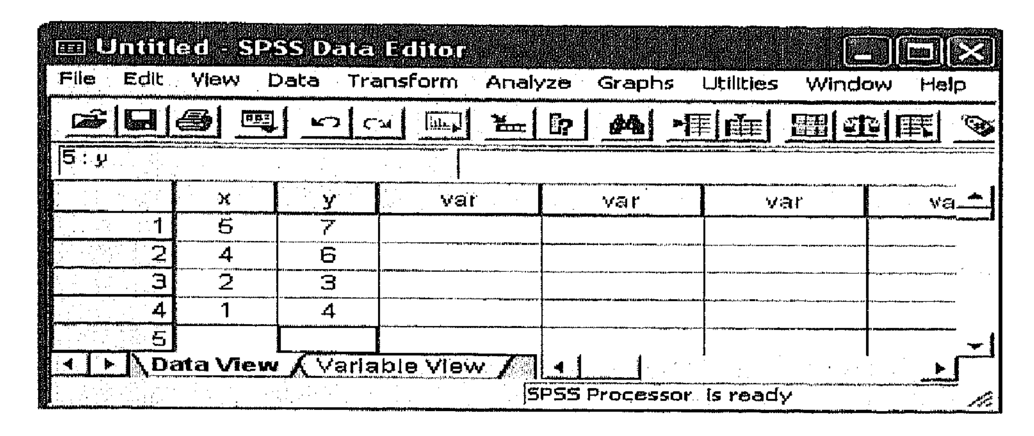


الخطوات المتبعه لتعريف المتغيرات: لنفرض أننا نريد تعريف المتغيرين x ، y

انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه



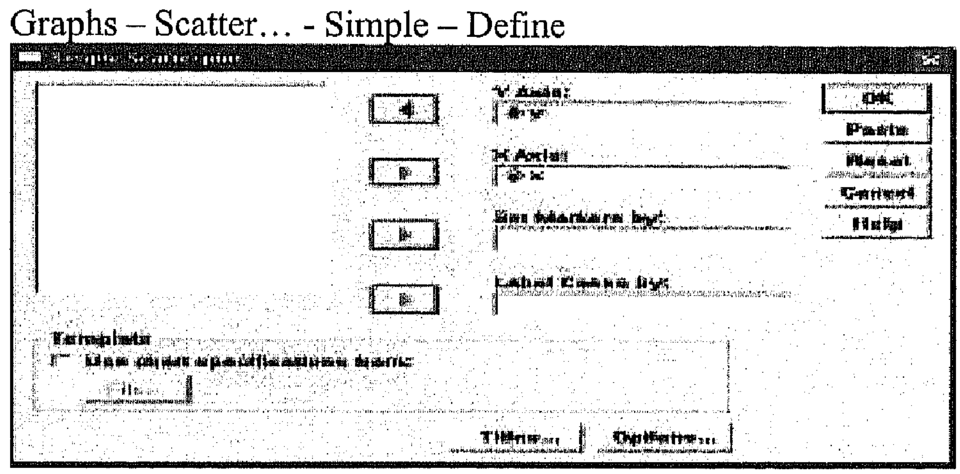
إدخال البيانات Data Input



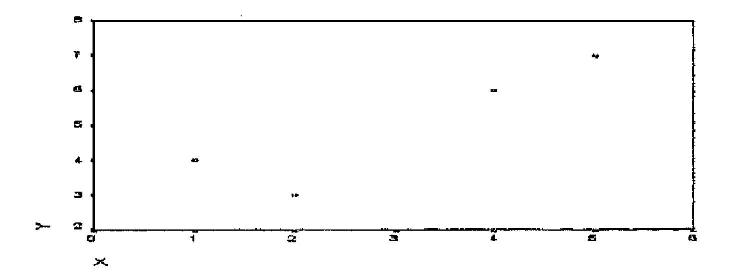
1- هل تعتقد بوجود علاقة بين العلامتين؟ ولماذا؟

للإجابة على هذا السؤال نقوم برسم شكل الانتشار

Graphs – Scatter… - Simple – Define



ثم نضغط زر OK فيظهر شكل الانتشار كما هو موضح أدناه:



شكل الانتشار يدل على وجود علاقة طردية

2- احسب معامل ارتباط بيرسون بين هذين المتغيرين؟

3- احسب معامل ارتباط سبيرمان للرتب بين هذين المتغيرين؟

Analyze – Correlate - Bivariate

اختر المتغيرين الكميين x, y وانقلهما الى نافذة :Variables

من Correlation Coefficients اختر Person و Spearman بالنقر على مربعهما

انقر زر Ok تظهر شاشة المخرجات أدناه:

Correlations

		X	Y
X	Pearson Correlation	1.000	.900
	Sig. (2-tailed)	.	.100
	N	4	4
Y	Pearson Correlation	.900	1.000
	Sig. (2-tailed)	.100	.
	N	4	4

			X	Y
Spearman's rho	X	Correlation Coefficient	1.000	.800
		Sig. (2-tailed)	.	.200
		N	4	4
	Y	Correlation Coefficient	.800	1.000
		Sig. (2-tailed)	.200	.
		N	4	4

مثال: حساب معاملات الارتباط الجزئية Partial Correlation

استخدم مفهوم العلاقات مع الآخرين

المتغيرات: العلاقات مع الزملاء Friends، العلاقات مع الطلاب Students، العلاقات مع المدراء Managers، العلاقات العامة General

سؤال الدراسة:

هـــل اعـــضاء هيئة التدريس الذين لديهم علاقات عالية في احد المجالات، يكون لديهم علاقات جيدة في الابعاد الأخرى اذا كان لديهم المستوى نفسه من العلاقات العامة.

* ادخال البيانات Input Data

	friends	students	managers	general
1	5	6	7	8
2	6	7	5	9
3	4	3	5	4
4	7	7	8	9
5	4	5	4	5
6	10	9	10	9
7	7	6	7	8
8	3	4	3	6
9	4	5	4	8
10	6	7	6	7

Analyze – Correlate – Bivariate…

Bivariate Correlations
Variables:
friends
students
managers
general
OK
Paste
Reset
Cancel
Help
Correlation Coefficients
Pearson Kendall's tau-b Spearman
Test of Significance
Two-tailed One-tailed
Flag significant correlations
Options...

انقل المتغيرات الى نافذة :Variables

انقر مربع الاختيار لمعامل ارتباط Person

انقر زر ...Options ثم انقر مربع اختيار Means and standard deviations

Bivariate Correlations: Options
Statistics
Means and standard deviations
Cross-product deviations and covariances
Missing Values
Exclude cases pairwise
Exclude cases listwise
Continue
Cancel
Help

انقر زر Continue، ثم انقر زر Ok تظهر شاشة المخرجات أدناه:

Correlations

Descriptive Statistics الاحصاءات الوصفية للمتغيرات التي تم إختيارها

	Mean	Std. Deviation	N
FRIENDS	5.60	2.066	10
STUDENTS	6.10	1.853	10
MANAGERS	5.90	2.132	10
GENERAL	6.50	1.509	10

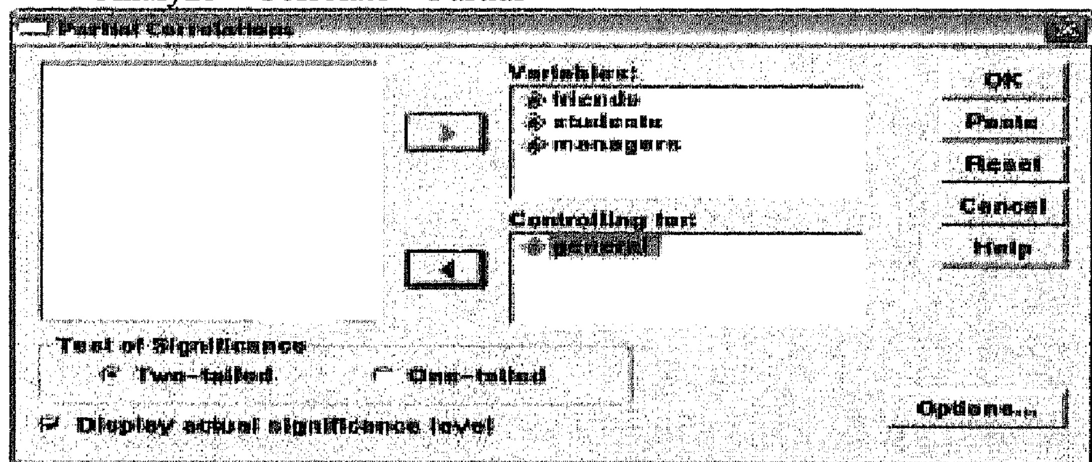
Correlations معاملات ارتباط بيرسون للمتغيرات التي تم اختيارها

		FRIENDS	STUDENTS	MANAGERS	GENERAL
FRIENDS	Pearson Correlation	1	.913**	.924**	.820**
	Sig. (2-tailed)	.	.000	.000	.004
	N	10	10	10	10
STUDENTS	Pearson Correlation	.913**	1	.790**	.934**
	Sig. (2-tailed)	.000	.	.006	.000
	N	10	10	10	10
MANAGERS	Pearson Correlation	.924**	.790**	1	.673*
	Sig. (2-tailed)	.000	.006	.	.033
	N	10	10	10	10
GENERAL	Pearson Correlation	.820**	.934**	.673*	1
	Sig. (2-tailed)	.004	.000	.033	.
	N	10	10	10	10

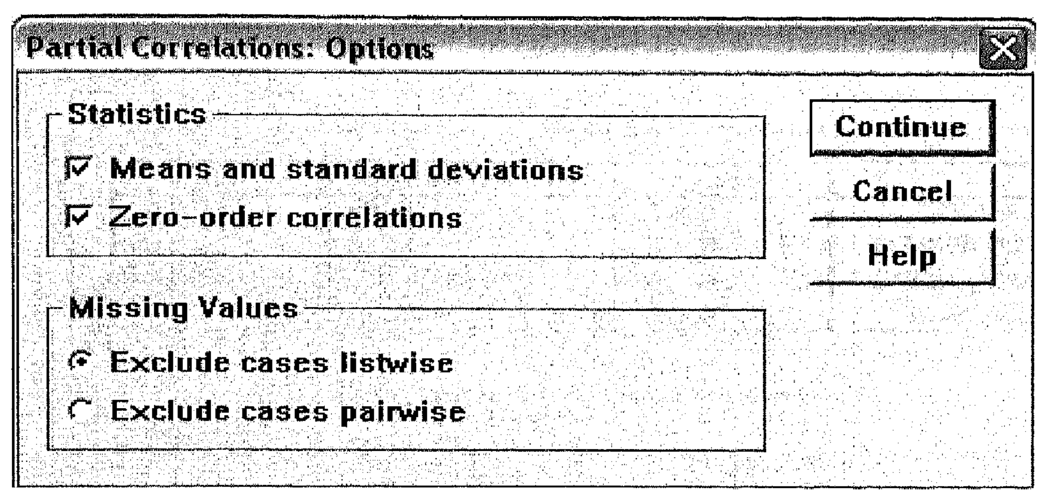
**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Analyze – Correlate – Partial



ضع المتغيرات friends, students, managers في نافذة:Variables وضع المتغير general في نافذة:Controlling for، انقر زر...Options تظهر الشاشة أدناه:



انقر مربعي Means and standard deviations و Zero-order correlations

انقر زر Continue ، ثم انقر زر Ok تظهر شاشة المخرجات أدناه:

➜ Partial Corr

Variable	Mean	Standard Dev	Cases
FRIENDS	5.6000	2.0656	10
STUDENTS	6.1000	1.8529	10
MANAGERS	5.9000	2.1318	10
GENERAL	6.5000	1.5092	10

المتوسطات Mean والانحرافات المعيارية Standard Dev وعدد الأفراد Cases

- - - P A R T I A L   C O R R E L A T I O N   C O E F F I C I E N T S - - -

Zero Order Partials

	FRIENDS	STUDENTS	MANAGERS	GENERAL
FRIENDS	1.0000	.9116	.9235	.8198
	(0)	(8)	(8)	(8)
	P= .	P= .000	P= .000	P= .004
STUDENTS	.9116	1.0000	.7904	.8337
	(8)	(0)	(8)	(8)
	P= .000	P= .	P= .006	P= .003
MANAGERS	.9235	.7904	1.0000	.6734
	(8)	(8)	(0)	(8)
	P= .000	P= .006	P= .	P= .033
GENERAL	.8198	.8337	.6734	1.0000
	(8)	(8)	(8)	(0)
	P= .004	P= .003	P= .033	P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

معاملات الارتباط الثنائية Zero-Order Correlations

```
- -PARTIAL  CORRELATION  COEFFICIENTS - -

Controlling for..  GENERAL

             FRIENDS    STUDENTS   MANAGERS

FRIENDS       1.0000      .7127      .8774
             (     0)    (     7)   (     7)
             P= .        P= .031    P= .002

STUDENTS       .7127     1.0000      .6107
             (     7)    (     0)   (     7)
             P= .031     P= .       P= .081

MANAGERS       .8774      .6107     1.0000
             (     7)    (     7)   (     0)
             P= .002     P= .081    P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed
```

معاملات الارتباط الجزئية Partial Correlations

النتيجة:

تم حساب المتوسطات Mean والانحرافات المعيارية Standard Dev وعدد الأفراد Cases لكل متغير من المتغيرات، كما حسبت معاملات الارتباط الثنائية Zero-Order Correlations فإذا كانت قيمة مستوى الدلالة اقل من المستوى المقبول 0.05 فإن معامل الارتباط يكون مقبولاً إحصائيا، وقد حسبت معاملات الارتباط الجزئية Partial Correlations فإذا كانت قيمة مستوى الدلالة اقل من المستوى المقبول 0.05 فإن معامل الارتباط الجزئي تكون مقبولة إحصائيا، أما إذا كانت قيمة مستوى الدلالة أكبر من المستوى المقبول 0.05 فإن قيمة معامل الارتباط الجزئي تكون غير مقبولة إحصائيا ويمكن إقرار عدم وجود علاقة بين المتغيرين.

وحتى نقلل من احتمال رفض الفرضية الصفرية وهي صحيحة (الخطأ من النوع الأول) يجب تعديل مستوى الدلالة ليصبح 0.05 مقسوماً على عدد معاملات الارتباط المحسوبة (3 في

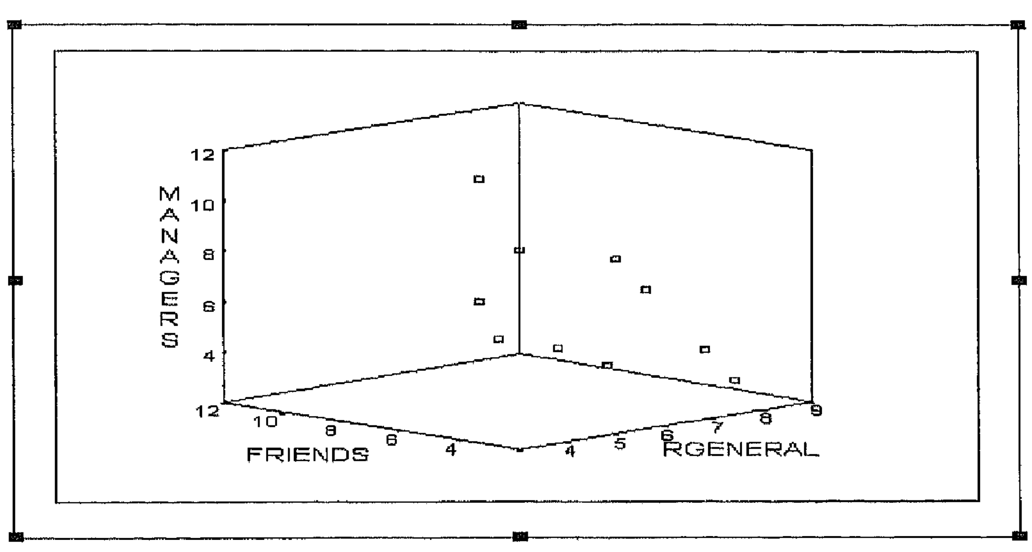
هذا المثال) لتصبح 0.0167=3/0.05، وباستخدام هذا المعيار فإن معاملات الارتباط الجزئية بين Friends و Manager هو الارتباط الجزئي المقبول إحصائيا من اصل الارتباطات المحسوبة.

استخدام الرسوم البيانية لتوضيح النتائج

لوحة الانتشار ثلاثية الأبعاد 3D-Scatterplot

Graph - Scatter… - 3-D – Define

انقر زر Ok تظهر شاشة المخرجات أدناه:



لوحة الانتشار ثلاثية الأبعاد 3D-Scatterplot

6-7 تمارين Exercise.

س1: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- حسب معامل الارتباط بين X , Y فكان 0.65- ، عرف المتغيران $L = 2X - 0.2$, $M=-3X+0.3$ فإن معامل الارتباط الجديد بين L , M هو:

أ- (0.35-)   ب- (0.85-)   ج- (0.65-)   د- 0.65

2- إذا كان $\Sigma(x-x')(y-y') = -27$ ، وكان $\Sigma(x-x')^2 = 36$ ، وكان $\Sigma(y-y')^2 = 81$ ، فإن معامل ارتباط بيرسون يساوي $r_{xy} =$?

أ- $\frac{(-27)}{36}$ ب- $\frac{(-27)}{81}$ ج- $\frac{(-27)}{9}$ د- $\frac{(-27)}{54}$

3- احدى القيم التالية يمثل معامل ارتباط سلبي عال جداً:

أ- (0.7-)   ب- (0)   ج- (0.98-)   د- (0.99)

4- إذا كانت قيم X كلها 5,5,5,5,5,... وقيم Y كلها -3،-3،-3،-3،3،... فإن معامل ارتباط سبيرمان=

أ- 1   ب- (1-)   ج- 0   د- 0.12

5- إذا أردنا قياس قوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الاسلوب الإحصائي المناسب هو:

أ- تحليل الارتباط   ب- تحليل الانحدار   ج- التوقع الرياضي   د- احتمال الحدث

6- إذا أعطيت بيانات عن متوسط درجات الحرارة وكمية الاستهلاك من الكهرباء خلال 50 يوم من أيام الصيف فإنه يمكن القول أن اتجاه العلاقة بين المتغيرين:

أ- عكسي   ب- طردي   ج- قوي   د- ضعيف

7- نوع الارتباط بين مساحة المربع وأبعاده هو:

أ- تام، بسيط ب- غير تام، بسيط ج- غير تام، متعدد د- تام، متعدد

8- عند حساب معامل الارتباط بين لون العيون ل (50) طالبة ونسبة الذكاء كان معامل الارتباط (0.9) فإنه يمكن تفسير هذا الارتباط القوي ل:

أ- التأثر بمتغير ثالث فقط ب- التأثر بأكثرمن متغير ج- العلاقةالتبادلية د-عامل الصدفة

9- القيمة التي تمثل أقوى معامل ارتباط عكسي مما يلي هو:

أ- (-2) ب- (-0.7) ج- (-0.90) د- (-1.5)

10- قام طالب بحساب معامل ارتباط بيرسون بين متغيرين فوجد أنه = (1.3) فإن ذلك يدل على:

أ- خطأ في الحساب ب- عدم وجود ارتباط

ج- طردي تام د- عكسي تام

11- إذا كان مجموع مربعات فروق الرتب بين (6 قيم) للمتغيرين X, Y هو (50)، فإن معامل ارتباط سبيرمان للرتب بين X, Y يساوي:

أ- (-1.43) ب- (-0.43) ج- (-0.57) د- (0.43)

12- إذا كان $\Sigma(x-x')\ (y-y') = -5$ ، وكان $\Sigma(x-x')^2 = 9$ ، وكان $\Sigma(y-y')^2 = 4$ ، فإن معامل ارتباط بيرسون يساوي $r_{xy} =$?

أ- $\frac{(-5)}{36}$ ب- $\frac{(-5)}{4}$ ج- $\frac{(-5)}{6}$ د- $\frac{(-5)}{9}$

13-حسب معامل ارتباط سبيرمان بين X, Y فكان 0.6 فإذا كانت n=10 ، فإن $\Sigma\ d^2$ يساوي

أ- 44 ب- 66 ج- 0.4 د- 0.66

14- إذا كانت r1=0.6 , r2=0, r3=0.43 , r4= -0.9 فإن معامل الارتباط الذي يعبر عن أقوى علاقة هو:

أ- r1 ب- r2 ج- r3 د- r4

س2: البيانات في الجدول أدناه هي لعينة من نزلاء الفندق حجمها n=10 تخص فترة إقامة كل نزيل في الفندق بالأيام Xi ومعدل الأجور في اليوم بالدينار Y والمطلوب:

رقم النزيل	الإقامة بالأيام X	تكلفة اليوم بالدينار Y
1	100	6
2	80	10
3	30	15
4	10	20
5	90	12
6	60	11
7	20	30
8	95	5
9	15	25
10	50	8

1. رسم شكل الانتشار بين X,Y
2. إيجاد معامل الارتباط البسيط r بين فترة الإقامة Xi ومعدل الأجور بالدينار Y

س3: الجـــدول التالي يضم بيانات تخص أعمار عينة من الموظفين X1 والخبرة الوظيفية بالسنين X2 والراتب الشهري بالدينار Y والمطلوب:

مسلسل	العمر بالسنوات X1	سنوات الخبرة X2	الراتب الشهري Y
1	25	2	150
2	30	4	170
3	32	5	190
4	35	7	200
5	37	8	220
6	40	10	250
7	45	12	300
8	50	15	350
9	55	18	370
10	60	20	500

1. إيجاد معامل الارتباط البسيط r بين سنوات الخبرة X2 والراتب الشهري Y
2. إيجاد معامل الارتباط المتعدد R
3. إيجاد معامل ارتباط الجزئي $r_{y2.1}$ و $r_{21.y}$ مع تفسير النتائج.

س4: اجري بحث عن العلاقة بين طول لاعب كرة السلة والقدرة على التسجيل لدى عينة من اللاعبين، والباحث يفترض أن العلاقة هي نتيجة اللياقة البدنية، بمعنى أن اللذين يتدربون أكثر تصبح لديهم قدرة اكبر، ولفحص الفرضية تم تدوين بين طول اللاعب وقدرته على التسجيل لدى 10 لاعبين، كما قام بتدوين عدد الساعات الأسبوعية التي يستغرقها اللاعب في التمرين، إذا أراد الباحث فحص العلاقة بين طول اللاعب والقدرة على التسجيل بعد استبعاد أثر عدد ساعات التمرين (بافتراض أن جميع اللاعبين يتدربون العدد نفسه من الساعات.

عدد ساعات التمرين	القدرة على التسجيل	طول اللاعب
20	20	180
20	18	185
20	22	190
20	17	200
20	15	188
20	20	189
20	25	201
20	16	195
20	20	197
20	22	202

الفصل الثامن

اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

الفصل الثامن

اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

1-8 اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

يستخدم لإصدار تعميمات تتعلق بالنسب، وتستخدم للظواهر التي لا يمكن قياسها ولكن يمكن تقدير نسبة وقوعها. وهنا لا بد ان نتذكر أن المتغيرات التي نتعامل معها هي متغيرات تصنيفية يعني انها تقع في المستوى الاسمي Nominal من مستويات القياس.

مثال: هل تختلف نسبة الموافقين على البحث التربوي باختلاف الجنس؟

مثال: هل تختلف نسبة المعاقين في الأردن عن نسبتهم في سوريا؟

في حالة المجتمع:

يرمز لنسبة وقوع الظاهرة في المجتمع بالرمز P وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في المجتمع بـ Q وهي نسبة الفشل، وهي (1-P).

أما في حالة العينة:

يرمز لنسبة وقوع الظاهرة في العينة بالرمز p' وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في العينة بالرمز P وهي نسبة الفشل.

2-8 اختبار الفرضيات حول نسبة واحدة

Single Proportion (P)

- اذا كانت الفرضية غير متجهه

نسبة الحاصين على شهادات عليا في المجتمع الأردني.

نسبة وحدات الإنتاج الصالحة في مصنع معين.

يؤول هذا التوزيع إلى التوزيع الطبيعي عندما تصبح n>=5 وكذلك n(n-1)>=0

ويكون الخطأ المعياري للتوزيع $\sigma L = \sqrt{n(1-n)}$

$$Z = \frac{p' - P}{\sqrt{PQ/n}}$$

حيث أن: n : حجم العينة

نسبة وقوع الظاهرة في العينة ب 'p وهي نسبة النجاح.

نسبة وقوع الظاهرة في المجتمع ب P وهي نسبة النجاح.

نسبة عدم وقوع الظاهرة في المجتمع ب Q وهي نسبة الفشل، وهي (1-P).

مثال8-2: يدعي احد المصانع بأن ما لا يقل عن 94% من إنتاجه مطابق للمواصفات المطلوبة، وللتأكد من صحة هذا الادعاء أخذت عينة من إنتاج المصنع حجمها 100 وحدة وبعد فحصها وجد بأن 90% منها كانت مطابقة للمواصفات. فهل تدعم هذه النتيجة ادعاء المصنع بمستوى دلالة = 0.05

$$H_0: P \geq 0.94$$
$$H_1: P < 0.94$$

القيمة الحرجة $Z_{0.05} = 1.64$

$$Z = \frac{p' - P}{\sqrt{PQ/n}} = \frac{0.90 - 0.94}{\sqrt{(0.94)(0.06)/100}} = \frac{-0.04}{0.02234} = -1.7899 = -1.79$$

النتيجة: بما أن القيمة Z المحسوبة (-1.79) أقل من القيمة الحرجة 1.64 نقبل الفرضية الصفرية ونستنتج بأن ادعاء صاحب المصنع صحيح، وانه ليس هناك فرق جوهري بين نسبة المجتمع P ونسبة العينة 'p عند مستوى الدلالة $\alpha=0.05$

مثال8-2: يتوفر في الأسواق دواء وعلى أساس نسبة نجاحه في تخفيض توتر الأعصاب هي 60% وظهر دواء جديد لنفس المرض كان قد تم تجربته على عينة تتكون من 100 شخص ودلت النتائج على شفاء 70 شخصاً منهم باستخدام هذا الدواء الجديد. فهل يمكن الاستنتاج من أن الدواء الجديد هو أفضل من النوع المتوفر في الأسواق عند مستوى معنوية 0.05

الحل:

$$H_0: P = 0.6$$
$$H_1: P > 0.6$$

قيمة Z الجدولية عند مستوى الدلالة $\alpha=0.05$ الاختبار بذيل واحد، $Z_{0.05} = 1.64$

قيمة Z المحسوبة:

$$Z = \frac{p' - P}{\sqrt{PQ/n}} = \frac{0.7 - 0.6}{\sqrt{(0.6)(0.4)/100}} = \frac{0.1}{0.0489} = 2.044 = 2.04$$

النتيجة: بما أن القيمة Z المحسوبة (2.04) أكبر من القيمة الحرجة ($Z_{0.05} = 1.64$) نرفض الفرضية الصفرية H0 القائلة أن نسبة نجاح الدواء الجديد هي 60% ونستنتج بأن نسبة نجاحه أكثر من 60% أي أنه أفضل من الدواء المستخدم في الأسواق.

- اذا كانت الفرضية متجهه

$$Z = \frac{L - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

مثال 8-3: إذا كان من المتوقع أن تكون نسبة الإناث في الدولة المتقدمة من بين ذوي الأعمار التي تزيد عن 50 سنة تساوي 0.5 ورغب باحث في معرفة إن كانت هذه النسبة تزيد عن ذلك في الأردن.فاختارعينة عشوائية طبقية حسب المستوى الاقتصادي الاجتماعي ممن تزيد أعمارهم عن 50 سنة. وكان حجم العينة 400 شخص فوجد أن عدد الإناث في هذه العينة 242 فهل تدعم هذه البيانات ما يدعيه الباحث؟

$H0: \pi = 0.5$

$H1: \pi > 0.5$

$\alpha = 0.05 \quad Z = 1.64$ القيمة الحرجة

$\alpha = 0.01 \quad Z = 2.34$ القيمة الحرجة

$L = 242/400 = 0.605$ النسبة

القيمة المحسوبة

$$Z = \frac{L - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

$$Z = \frac{0.605 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{400}}}$$

$$Z = \frac{0.105}{\sqrt{\frac{0.25}{400}}} = \frac{0.105}{0.025} = 4.2$$

بمـــا أن قيمة Z المحسوبة (4.2) > الحرجة (1.64، 2.34) على مستوى $\alpha = 0.01$ يمكن أن نقول أن نسبة الإناث في الأردن من بين الذين تتجاوز أعمارهن 50 سنة تزيد بدلالة إحصائية مرتفعة ($P<0.01$) عن النسبة المناظرة لها في الدول المتقدمة.

3-8 اختبار الفرضيات حول نسبتين مستقلتين

تكون النسبتان مستقلتين إذا كانتا تعودان إلى مجتمعين مستقلين.

مثال:

مقارنة نسبة المقبولين من الذكور لوظيفة معينة إلى نسبة المقبولات من الإناث.

مقارنة نسبة المصابين بالسرطان من الذكور إلى نسبة المصابات من الإناث.

مقارنة نسبة المقبولين من الذكور في الجامعة إلى نسبة المقبولات من الإناث.

$H_0: \pi_1 = \pi_2$

π_1 : النسبة في المجتمع الأول. π_2 : النسبة في المجتمع الثاني.

$\hat{\sigma}_L$: التقدير للخطأ المعياري لتوزيع المعاينة.

$$\hat{\sigma}_L = \sqrt{\pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$\hat{\pi} = \frac{F1 + F2}{n1 + n2}$$

F1 : التكرار للظاهرة في العينة الأولى. F2 : التكرار للظاهرة في العينة الثانية.

n1 : عدد الأفراد في العينة الأولى. n2 : عدد الأفراد في العينة الثانية.

$$\mathbf{Z = \frac{L1 - L2}{\sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

مثال8-4: جرى تطبيق طريقتين لتعليم مفهوم التسارع على مجموعتين عدد الأفراد في كل منها 100 وفي نهاية التدريس وجد بأن عدد الذين أتقنوا المفهوم في المجموعة الأولى =68 وأن العـــدد المناظـــر في المجموعة الثانية =54 ، فهل تختلف الطريقتان جوهرياً في نسب الذين يتقنون المفهوم في كل منها؟

$H_0: \pi_1 = \pi_2$
$H_1: \pi_1 \neq \pi_2$

القيمة الحرجة على مستوى $\alpha=0.05$ هي 1.69 وعلى مستوى $\alpha=0.01$ هي 2.58

القيمة المحسوبة:

$$L1 = 68/100 = 0.68$$
$$L2 = 54/100 = 0.54$$
$$\hat{\pi} = \frac{F1 + F2}{n1 + n2} = \frac{68 + 54}{100 + 100} = 0.61$$
$$Z = \frac{L1 - L2}{\sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$
$$Z = \frac{0.68 - 0.54}{\sqrt{0.61(1-0.61)\left(\frac{1}{68} + \frac{1}{54}\right)}} = 2.03$$

النتيجة: بما أن Z المحسوبة (2.03) تقع في منطقة الرفض للفرضية الصفرية على مستوى $\alpha=0.05$ يمكن القول أن نسبتي متقني مفهوم التسارع من المجموعتين تختلفان اختلافاً ذا دلالة إحصائية $(P<0.05)$.

P_1 : نسبة النجاح في المجتمع الأول. P_2 : نسبة النجاح في المجتمع الثاني.

p'_1 : نسبة النجاح في العينة الأول. p'_2 : نسبة النجاح في العينة الثانية.

$$\mu_{p'1} = P_1$$
$$\sigma_{p'1} = \sqrt{\frac{P_1Q_1}{n1}}$$

مثال 8-5: في استطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج تربوي معين كان عدد الموافقين من المدرسين 33 ومن المدرسات 18، هل تختلف النسبة في المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- تختلف وذلك لأن $Z=1.177$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا تختلف وذلك لأن $Z=1.177$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

$H_0: \pi_1 = \pi_2 \qquad H_1: \pi_1 \neq \pi_2$

القيمة الحرجة على مستوى $\alpha=0.05$ هي 1.69 وعلى مستوى $\alpha=0.01$ هي 2.58

القيمة المحسوبة:

$$L1 = 33/100 = 0.33$$

$$L2 = 18/100 = 0.18$$

$$\pi\hat{} = \frac{F1 + F2}{n1 + n2} = \frac{33 + 18}{100 + 100} = 0.255$$

$$Z = \frac{L1 - L2}{\sqrt{\pi\hat{}\,(1-\pi\hat{})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$Z = \frac{0.33 - 0.18}{\sqrt{0.255\,(1-0.255)\left(\frac{1}{33} + \frac{1}{18}\right)}} = 1.177$$

النتيجة:

بمـــا أن Z المحـــسوبة (1.177) تقـــع في منطقة القبول للفرضية الصفرية على مستوى $\alpha=0.05$ يمكن القول أن نسبتي متقني مفهوم التسارع من المجموعتين لا تختلفان اختلافاً ذا دلالة إحصائية ($P<0.05$).

4-8 اختبار الفرضيات حول نسبتين للبيانات الغير مستقلة (المترابطة)

تكون النسبتان غير مستقلتان إذا كانتا تعودان إلى مجتمعين غير مستقلين.

مثال: نسبة الموافقين على التعليم المختلط في الجامعات قبل استماعهم لمحاضرة تبين رأي الدين في ذلك وبعد استماعهم للمحاضرة.

$$H_0: \pi_1 = \pi_2$$

يستخدم اختبار χ^2 حيث تفرغ الإجابات في جدول التصاحب Contingency table حيث ترصد التكرارات داخل الجدول.

	بعد		
	نوع القرار	معارض	موافق
قبل	موافق	a	b
	معارض	c	d

القيمة المحسوبة

$$\chi_1^2 = \frac{(d - a)^2}{d + a}$$

نستخرج القيمة الحرجة للرفض أو القبول باستخدام جدول χ^2 بدرجة حرية واحدة وحسب مستوى الدلالة المطلوب.

ترفض الفرضية الصفرية إذا كانت قيمة χ^2 المحسوبة > أكبر من القيمة الحرجة.

تقبل الفرضية الصفرية إذا كانت قيمة χ^2 المحسوبة < أقل من القيمة الحرجة.

مثال8-5: تم استطلاع أراء 60 من التربويين حول مسألتين الأولى تتعلق بتأنيث التعليم في الصفوف الابتدائية الثلاثة الأولى والثانية تتعلق بتأنيث التعليم في المرحلة الابتدائية بأكملها. فإذا وافق 42 منهم على المسألة الأولى ووافق 34 منهم على المسألة الثانية. وكان جميع الذين وافقوا على المسألة الثانية قد وافقوا على المسألة الأولى. فهل تدل هذه البيانات على وجود اختلاف جوهري في نسب الموافقين على المسألتين؟

$H_0: \pi_1 = \pi_2$

$H_1: \pi_1 \neq \pi_2$

القيم الحرجة للمختبر الإحصائي χ^2 بدرجة حرية واحدة وحسب مستوى الدلالة $\alpha=0.05$ هي $\chi^2_{(1,0.95)} = 3.84$ ، وعلى مستوى $\alpha=0.01$ هي $\chi^2_{(1,0.99)} = 6.64$

القيمة المحسوبة:

	المسألة الثانية			
	نوع القرار	معارض	موافق	
المسألة الأولى	موافق	8	34	42
	معارض	18	0	18
		26	34	60

$$\chi^2 = \frac{(d-a)^2}{d+a} = \frac{(0-8)^2}{0+8} = \frac{64}{8} = 8$$

النتيجة:

بما أن القيمة المحسوبة ($\chi^2 = 8$) أكبر من القيمة الحرجة ($\chi^2_{(1,0.99)} = 6.64$) وعلى مستوى $\alpha=0.01$ نستطيع القول بأن نسبة التربويين الموافقين على تأنيث التعليم في الصفوف

الــثلاثة الأولى تختلف جوهرياً (P<0.01) عن نسبة الموافقين على مثل هذا النوع من التعليم وذلك عندما يتم استطلاع أرائهم على المسألتين في نفس الوقت.

5-8 اختبار χ^2 لحسن المطابقة
The Chi-Square Goodness of fit

يستخدم للمقارنة بين توزيع العينة والتوزيع النظري للمجتمع الذي سحبت منه.

اختــبار كــاي تربيع لحسن المطابقة، والغرض منه تحديد ما إذا كانت النسب الملاحظة تختلف جوهريا عن قيم متوقعة أو نظيرة لها . وهو مكافيء لأختبار Z للنسب.

ويحسب عن طريق المعادلة $\chi^2 = \Sigma(O_i-E_i)^2/E$

حيث Oi: التكرارات المشاهدة Ei : التكرارات المتوقعة

مثال: اجرى معلم دراسة لمعرفة نسب الرسوب في المواد التي يدرسها والتي على اساسها يتم تقسيم الطلاب الى مستويات وقد فام بدراسة شملت 400 طالب كما يلي:

المستويات	1	2	3	4	5
عدد الطلاب	900	700	800	1100	700
نسبة الرسوب	%20	%15	%10	%30	%25

اختبر χ^2 لمعرفة مدى تطابق التوزيع.

الحل: نحسب التكرارات المتوقعة كما يلي

المستوى	التكرارات المشاهدة	التكرار المتوقع
1	900	800=0.20*4000
2	700	600=0.15*4000
3	800	400=0.10*4000
4	1100	1200=0.30*4000
5	700	1000=0.25*4000
المجموع	4000	3000

ايجاد قيمة χ^2 المحسوبة من المعادلة $\chi^2 = \Sigma(O_i-E_i)^2/E$

$$\chi^2= (900-800)^2/800+(700-600)^2/600+(600-400)^2/400+(1100-1200)^2/1200+(700-000)^2/1000$$

$$\chi^2 = 12.5 + 16.67 + 100 + 8.3 + 9 = 146.47$$

قيمة χ^2 الجدولية $\chi^2_{(4, 0.01)} = 13.28$

القرار:

بمـــا ان قيمة χ^2 الجدولية اقل من المحسوبة نرفض الفرضية الصفرية ونقبل البديلة، وهذا يعني ان توزيع العينة يختلف عن التوزيع النظري للمجتمع الذي سحبت منه.

2- اختبار χ^2 للاستقلالية The Chi-Square Test of Independence

لمعرفة اذا كان متغيرين مستقلين عن بعضهما.

يـــستخدم لمعـــرفة ما إذا كانت الآراء حول قضية معينة تختلف أو ترتبط مع متغير اسمي آخر، وعند استخدام هذا النوع من الاختبار فإنه يتم إنشاء جدول التوافق.

6-8 استخدام برمجية SPSS في حل المسائل.

1- اختبار كاي تربيع لحسن المطابقة، والغرض منه تحديد ما إذا كان النسب الملاحظة تختلف جوهريا عن قيم متوقعة أو نظيرة لها .

خطوات استخراجه من SPSS :

- ادخل البيانات تصاعدياً
- من قائمة Analyze أختر Nonparametric Tests
- اختر Square–Chi
- انقل المتغير إلى خانة Variable List
- من خانة Expected Value أشر على مربع Value
- ادخل النسب حسب التكرارات في محرر البيانات
- قم بوزن الحالات على أساس التكرارات من قائمة Weight Cases Data

مـــثال: إذا كانـــت نسبة الذكاء بين طلاب الثانوية العامة تتوزع بين ثلاث مستويات (2: 7: 1) على التوالي مرتفعي الذكاء ومتوسطي الذكاء ومحدودي الذكاء، وأراد باحث أن يختبر عينة ممثلة من مدرسة محددة لهذا المجتمع بحجم 400 طالب فوجد أن عدد عالي الذكاء = 50 ومتوســـط الـــذكاء = 290 ومحدودي الذكاء = 60 فهل يمكن اعتبار هذه العينة ممثلة للمجتمع.

2- اختـــبار كاي تربيع للاستقلالية: يستخدم لمعرفة ما إذا كانت الآراء حول قضية معيـــنة تخـــتلف أو ترتبط مع متغير اسمي آخر، وعند استخدام هذا النوع من الاختبار فإنه يتم إنشاء جدول التوافق.

خطوات استخراجه من SPSS :

- من قائمة Analyze أختر Descriptive Statistics ثم Crosstabs
- أنقل المتغيرات المطلوبة إلى الصفوف والأعمدة
- أنقر مربع Statistics ثم اختر Chi - Square
- أنقر مربع Cells وهذا يمكنك من إجراء إحصائيات إضافية
- قم بوزن الحالات على أساس التكرارات من قائمة Weight Cases Data

مثال: قام مدير التعليم بأحد مناطق المملكة بسؤال 200 من التربويين حول الاستمرار في الـــدوام الـــصيفي أم لا ، وسأل ثلاث فئات (مشرفين، معلمين، ومدراء المدارس) وكانت الإجابات على النحو التالي (المشرفين 38 موافق و12 معارض، المعلمين 26 موافق و 102 معـــارض، المد راء 16 موافق و 6 معارض) فهل تدل البيانات على اختلاف الرأي حول هذه المسألة باختلاف الوظيفة؟

7-8 تمارين Exercise.

تمرين (1) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- اذا كانت n=100 فان القيمة العظمى لخطأ المعاينة المعطى بالعلاقة $\sigma_L = \sqrt{\frac{\pi(1-\pi)}{n}}$ ستكون عند القيمة التالية للمعلم π :

أ- 0.5 ب- 0.9 ج- صفر د- 0.1

2- أي من الرموز التالية يشير إلى القيمة الحرجة لاختبار χ^2 عندما يكون مستوى الدلالة $\alpha=0.01$ وتكون درجة الحرية $\gamma = 2$

أ- $\chi^2_{2,0.01}$ ب- $\chi^2_{0.99}$ ج- $\chi^2_{2,0.99}$ د- $\chi^2_{2,0.90}$

3- أي من التالي يكون هو الأكبر؟

أ- $\chi^2_{1,0.01}$ ب- $\chi^2_{1,0.99}$ ج- $\chi^2_{1,0.99}$ د- $\chi^2_{1,0.90}$

4- أي من التالي يكون هو الأكبر؟

أ- $\chi^2_{1,0.95}$ ب- $\chi^2_{2,0.95}$ ج- $\chi^2_{3,0.95}$ د- $\chi^2_{1,0.90}$

5- اختبار يتألف من 40 فقرة حصل طالب على 26 فقرة بشكل صحيح هل يختلف هذا العدد من الفقرات عما هو متوقع بالتخمين؟ إذا كانت الفقرات من نوع صح/خطأ.

أ- لا، يختلف لأن Z=0.897 و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ب- نعم، يختلف لأن Z=0.897 و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ج- لا، يختلف لأن Z=0.897 و $\chi^2 = 3.6$ وهي ذات دلالة احصائية.

د- نعم، يختلف لأن Z=0.897 و $\chi^2 = 3.6$ وهي ذات دلالة احصائية.

6- اختبار يتألف من 40 فقرة حصل طالب على 26 فقرة بشكل صحيح هل يختلف هذا العدد من الفقرات عما هو متوقع بالتخمين؟ إذا كان لكل فقرة 4 بدائل.

أ- لا، يختلف لأن Z=0.897 و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ب- نعم،تخـــتلف لأن Z=5.84 و34.13= χ^2 هي ذات دلالة احصائية على مستوى $\alpha=0.001$

ج- لا، يختلف لأن Z=0.897 و 3.6 = χ^2 وهي ذات دلالة احصائية.

د- لا،تخـــتلف لأن Z=5.84 و34.13= χ^2 هـــي ذات دلالة احصائية على مستوى $\alpha=0.001$

7- في اســـتطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج تربوي معين كان عـــدد الموافقين من المدرسين 33 ومن المدرسات 18، هل تختلف النسبة في المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- تختلف وذلك لأن Z=2.433 وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا تختلف وذلك لأن Z=2.433 وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

8- في اســـتطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج تربوي معين كان عـــدد الموافقين من المدرسين 33 ومن المدرسات 18، هل تختلف النسبة في المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- تختلف وذلك لأن Z=2.433 وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا تختلف وذلك لأن Z=2.433 وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا تختلف وذلك لأن $\chi^2{}_2 = 5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

الفصل التاسع

تحليل التباين

Analysis of Variance

9-1 مقدمة Introduction
9-2 تحليل التباين الأحادي One-Way Analysis of Variance
9-3 تحليل التباين الثنائي Tow-Way Analysis of Variance
9-4 تحليل التغاير Analysis of Covariance
9-5 استخدام برمجية SPSS في حل المسائل.
9-6 تمارين Exercise

الفصل التاسع

تحليل التباين
Analysis of Variance

1-9 مقدمة Introduction

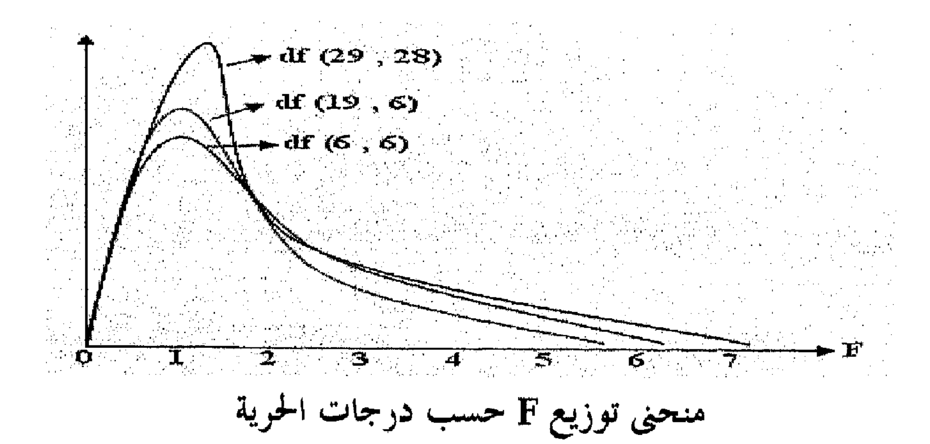
تحلـــيل التـــباين **ANalysis Of VAriance (ANOVA)**: هـــو تعميم لاختبار t ويـــستخدم لاختـــبار أكثر من عينتين، ويستخدم لمقارنة عدة متوسطات حسابية في آن واحد $H_0: \mu_1 = \mu_2 = \ldots = \mu_k$

تحلـيل التباين هو طريقة ذكية لاختبار اختلاف أوساط أكثر من مجموعتين دفعة واحدة من خلال التباين.

يمكـــن استخدام الرسم البياني Box Plot لتوضيح نتائج المقارنة بين متوسط أكثر من عينتين من العينات المستقلة.

توزيـــع F عبارة عن مجموعة من المنحنيات التكرارية يتميز كل منها عن الآخر برقمين لدرجات الحرية أحدهما يمثل درجة حرية للبسط والآخر درجة حرية للمقام . وقيمة F هي قيمة توضـــح نـــسبة التباين Variance ratio لعينتين والرمز F إشارة إلى العالم Fisher الذي قام بعمـــل هذا الاختبار والمعروف باختبار F . وقد قام العالم Snedecor بحساب جداول خاصة لـــتوزيع F وفـــيها درجات الحرية التي في أعلى الجدول تخص البسط أما درجات الحرية على العمود الجانبي فتخص المقام.

توزيع F : هو توزيع ملتو جهة اليمين بمعلمتين تتمثلان بدرجتي حرية (البسط ، المقام) وهمـــا $k - 1$ للبسط ، $k - n$ للمقام حيث n مجموع إحجام العينات، فإذا كان لدينا اختبار لقياس معنوية الفرق بين التقديرين (F) نوجد F_α حيث α مستوى المعنوية المستخدم للفرضية H_0 الـــتي ترفض إذا كان $F > F_\alpha$ وإلا نؤكد بوجود الاختلاف بين المتوسطات، والشكل التالي يبين توزيع F.



منحنى توزيع F حسب درجات الحرية

وتحليل التباين هو عملية يقصد بها تقسيم مجموع مربعات الانحرافات عن المتوسط الحسابي إلى مكوناته أي إرجاع كل من هذه المكونات إلى مسبباتها.

وطريقة تحليل التباين تفيد في مقارنة عدد من المعاملات يزيد عن أثنين.

كما تمتاز طريقة تحليل التباين بأنه يمكن فيها استعمال كل البيانات المأخوذة من التجربة في حساب قيمة واحدة للانحراف القياسي يمكن بها مقارنة المجموعات أو المعاملات التجريبية.

وكلما زاد عدد المتوسطات كلما زاد احتمال الخطأ وقل احتمال اتخاذ قرار صحيح ففي المثال الخاص بمقارنة متوسطات في ست مناطق فإنه يلزم إجراء الاختبار 15 مرة وبالتالي سوف ينخفض احتمال اتخاذ قرار صحيح في الخمسة عشر اختبار معاً من 0.95 إلى $(0.95)^{15}$ أي إلى 0.46 فقط وبالتالي يرتفع احتمال الخطأ في اتخاذ القرار الصحيح من مجرد 0.05 إلى 0.54 والذي يساوي(1-0.46) وهو احتمال كبير جداً للخطأ في اتخاذ القرار.

لذلك كان لابد من التفكير في أسلوب آخر بديل يوفر الوقت والمجهود وفي الوقت نفسه لا يقلل احتمال اتخاذ القرار الصحيح أو يكبر احتمال الخطأ في اتخاذ القرار، هذا الأسلوب هو الذي يسمى " تحليل التباين " والذي يختبر ما إذا كانت المتوسطات كلها متساوية مرة واحدة دون أخذهم اثنين اثنين ودون أن ينخفض احتمال اتخاذ قرار صحيح أو يزيد احتمال الخطأ عند اتخاذه. وهو الذي يسمى اختصاراً ANOVA وهو اختصار للمصطلح الإنجليزي Analysis of Variance.

ويعتمد هذا الأسلوب من أساليب التحليل الإحصائي على ما يعرف باختبار F والذي يعتمد أساساً على تحليل التباين. أن التباين ما هو إلا متوسط مربعات انحرافات القيم عن

وسطها الحسابي. أي أن التباين يعتمد أساساً على مجموع مربعات ثم القسمة على عدد المشاهدات. ويعتمد أسلوب تحليل التباين على تقسيم مجموع المربعات الكلي إلى أقسام فيمثل كل منهما أو يقيس أحد مصادر التغير أو الاختلاف Source of Variation يمثل أحدها – مثلاً – التغير بسبب المعاملات (أو المجتمعات) المختلفة ، ويمثل آخر التغير بسبب الأخطاء ثم تعرف الإحصائية (أو الاختبار) F بأنها خارج قسمة التباين بسبب المعاملات على التباين بسبب الأخطاء وهكذا. أي أنه يتم حساب التباين بسبب المعاملات، والتباين بسبب الأخطاء فيحصل على قيمة F المحسوبة وبمقارنة هذه القيمة بالقيمة الجدولية F نصل إلى قرار إما بقبول الفرضية الصفرية أو عدم قبولها عند مستوى المعنوية المطلوب. ولتحليل التباين تطبيقات كثيرة في مختلف المجالات.

إذا كان هناك متغير مستقل واحد، نسميه تحليل التباين الأحادي One-Way ANOVA

اذا كان هناك متغيرين مستقلين، نسميه تحليل التباين الثنائي Two-Way ANOVA

اذا كان هناك 3متغيرات مستقلة، نسميه تحليل التباين الثلاثي Three-Way ANOVA

ولكن يوجد متغير تابع وحيد في جميع تحليلات التباين أعلاه.

شروط استخدام اختبار تحليل التباين:

1. أن تكون العينات عشوائية مستقلة.
2. أن تكون العينات مسحوبة من مجتمعات لها توزيعات طبيعية.
3. أن تكون تباينات المجتمعات متساوية $\sigma^2_1 = \sigma^2_2 = \sigma^2_3 = \ldots\ldots = \sigma^2_k = \sigma^2$

ويتم إجراء اختبار تحليل التباين اعتماداً على اختبار F واعداد جدول تحليل التباين.

2-9 تحليل التباين الأحادي One-Way Analysis of Variance

إذا كان لكل فرد من أفراد العينة علامة على متغيرين، الأول يسمى المتغير العاملي Factor أو المتغير المستقل Independent Variable وهو متغير من النوع الأسمي Nominal أو الترتيبي Ordinal أما المتغير الآخر فهو المتغير التابع Dependent Variable وهو متغير من النوع الكمي وهو المتغير الذي سيتم فحص مساواة متوسطه لكل فئة من فئات المتغير المستقل.

إن قـيم x يتم تصنيفها الى k من العينات وفقاً لمعيار واحد معين، مثل درجات الطلاب في k من الشعب وكل شعبة تضم n من الطلاب.

يعتبر أبسط التصميمات التجريبية إذ فيه توزع معاملات التجربة Treatments علي كل الـوحدات التجـريبية أو الأفراد عشوائيا . بمعني أنه إذا كان لدينا في التجربة خمس معاملات مختلفة وكل من المعاملات سوف تنفذ من أربع وحدات فإن استعمال التوزيع العشوائي سوف يجعـل لكل مجموعة مكونة من أربع وحدات تجريبية فرصة متساوية لأن تعامل بأي معاملة من معاملات التجربة. أي أن كل وحدة تجريبية توضع في أي معاملة من المعاملات عشوائيا.

مميزات هذا التصميم :

1. يمكن استخدام أي عدد من المعاملات أو المكررات.
2. مفيد في ما إذا كانت الوحدات التجريبية قليلة أو من الصعب التحكم فيها كما في التجارب علي الحيوانات.
3. طـريقة التحليل بسيطة وسهلة حتى بفرض عدم تساوي المكررات (الوحدات) بين كل المعاملات.
4. تبقي طريقة التحليل الإحصائي بسيطة إذا ما فقدت قيم بعض الوحدات التجريبية أو حتى القيم الكلية لبعض المعاملات.

تحليل التباين الاحادي في حالة تساوي حجوم العينات

اذا كـان لديـنا عيـنات عشوائية حجم كل منها n ، مسحوبة من مجتمعات توزيعها طبيعي، ومتوسطاتها $\mu_1, \mu_2, \ldots\ldots, \mu_k$، وتباين σ^2، والمطلوب اختبار فرضية:

$H_0: \mu_1 = \mu_2 = \ldots\ldots = \mu_k$

$H_1: \mu_1 \neq \mu_2 \neq \ldots\ldots \neq \mu_k$

لاختـبار مساواة متوسطات المجموعات فإن تحليل التباين يستهدف تجزئة التباين الكلي للمتغير التابع إلى جزئين، ومن ثم تتم المقارنة بين تبايني الجزئين باستخدام اختبار F لذلك عملياً يجـزيء مجموع مربعات التباين ودرجات الحرية V الى تباين الأول معروف المصدر وهو بين المجمـوعات (Between Groups) ومصدرها الفروقات بين متوسطات المجموعات، فإذا كان

هـــذا الجـــزء كـــبيراً فإن متوسطات المحموعات غير متساوية، والثانية تباين ضمن المجموعات (Within Groups) وهـــي الجزء غير المعروف مصدره ويمكن أن يسمى الباقي (Residuals) أو الخطأ (Error).

$$SST = SSB + SSW$$

إن صيغة الاحصاءة المستخدمة لاختبار الفرضية هي:

$$F = \frac{MSB}{MSW} \quad .....F_{k-1,n-k}$$

القرار:

يكـــون القرار رفض الفرضية الصفرية H_0 التي تقول أن متوسطات المجموعات متساوية، اذا كانـــت نسبة التباين بين المجموعات الى التباين ضمن المجموعات كبيراً وهذه النسبة تسمى (قـــيمة F)، فـــإذا كانت قيمة F المحسوبة أكبر من القيمة الجدولية $F\alpha,\ k-1,\ k(n-1)$ نرفض الفرضية الصفرية التي تقول ان متوسطات المجموعات متساوية.

يعتـــبر هـــذا التصنيف هو أبسط أنواع تحليل التباين، حيث تصنف المشاهدات إلى عدة مجموعات على أساس متغير واحد أو خاصية واحدة.

والافتراضات الأساسية لهذا التحليل هي ما يلي:

1- نفترض أن عدد المجتمعات K وأنها جميعاً مستقلة.

2- أنها جميعاً تخضع للتوزيع الطبيعي بمتوسطات تساوي $\mu_1, \mu_2, ..., \mu_k$

3- أن لها جميعاً التباين نفسه σ^2 أي أن التباين لكل المجتمعات ثابت $= \sigma^2$

4- يتم سحب عينة عشوائية من كل من هذه المجتمعات وأن أحجام هذه العينات كلها متساوية وتساوي n ويمكن بكل بساطة افتراض عدم تساوي أحجام العينات ولن يختلف أسلوب التحليل على الإطلاق إلا في أشياء بسيطة جداً.

وتصنف البيانات عادة في هذا التحليل على النحو التالي :

المجتمع أو المعاملة	المشاهدات		
	1	2	n
1	y_{11}	y_{12}	y_{1n}
2	y_{21}	y_{22}	y_{2n}
:	:		
K	y_{k1}	y_{k2}	y_{kn}

حيث يمثل الصف الأول مشاهدات العينة الأولى أي المسحوبة من المجتمع الأول، ويمثل الصف الثاني مشاهدات العينة الثانية المسحوبة من المجتمع الثاني،.. وهكذا يمثل الصف الأخير مشاهدات العينة الاخيرة k المسحوبة من المجتمع الأخير رقم K.

كما تكون خطوات الاختبار كما يلي :

1- الفرض الصفري : هو أن متوسطات هذه المجتمعات متساوية، وبالرموز

$$HO : \mu_1 = \mu_2 = ... = \mu_k$$

2- الفرض البديل : هو أن بعض هذه المتوسطات غير متساوية (أو : يوجد متوسطان على الأقل غير متساويين).

3- إحصائية الاختبار : في هذه الحالة يرمز لها بالرمز F وتأخذ الشكل التالي :

$$F = \frac{S_R^2}{S_E^2}$$

والتي لها توزيع F بدرجات حرية للبسط K-1 وللمقام K(n-1) حيث : S_R^2 هو التباين بسبب الظاهرة أو المتغير أو المعاملات، S_E^2 هو التباين بسبب الخطأ.

ويمكن الحصول على الإحصائية F بتنظيم الحسابات في جدول يسمى " جدول تحليل التباين " ANOVA TABLE كما يلي :

جدول تحليل التباين (ANOVA)*

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	(الإحصائية) المحسوبة
بسبب المعاملات	SSR	K- 1	$S_R^2 = SSR/K-1$	$F = \frac{S_R^2}{S_E^2}$
بسبب الخطأ	SSE	K (n-1)	$S_E^2 = SSE / K(n-1)$	
الكلي	SST	N*k-1		

وسوف نوضح من المثال كيفية حساب المقادير الثلاثة SSE, SSR , SST

4- حدود منطقتي القبول والرفض: ويتم الحصول عليها من جدول توزيع F بدرجات حرية للبسط K-1 وللمقام ((k(n- 1). (اختبار الطرف الأيمن).

5- المقارنـــة والقـــرار: إذا كانت قيمة F المحسوبة من تحليل التباين أقل من قيمة F الجدولية نقبل الفرض العدمي بتساوي المتوسطات والعكس صحيح.

مـــثال 9-1 : البيانات التالية تمثل أعمار أربع عينات عشوائية من الناخبين سحبت من أربـــع مدن مستقلة (نفترض أن لها توزيعات طبيعية بمتوسطات $\mu4, \mu3, \mu2, \mu1$ وتباين مشترك يساوي σ^2) :

المدن k	المشاهدات						المجموع
	1	2	3	4	5	6	
الأولى 1	20	21	25	28	30	26	150
الثانية 2	23	22	27	20	26	20	138
الثالثة 3	19	20	21	28	20	18	126
الرابعة 4	24	29	30	28	27	24	162

والمطلـــوب اختـــبار الفـــرض العدمي بأن متوسطات أعمار الناخبين من المدن الأربع متساوية، أي أن المطلوب بالرموز هو :

$$HO : \mu1 = \mu2 = \mu3 = \mu4$$

وذلك بمستوى معنوية 5%

الحل: تكون خطوات الحل كما يلي:

1- الفرض الصفري:

$$HO : \mu1 = \mu2 = \mu3 = \mu4$$

2- الفــرض الــبديل: أن بعــض هــذه المتوسطات غير متساوٍ (اثنان على الأقل غير متساويين).

3- الإحصائية: وهي في هذه الحالة $F = \frac{S_R^2}{S_E^2}$ وتكون الحسابات التفصيلية لتحليل التباين كما يلي :

حيث K= 4   n= 6

أ) متوسطات الصفوف (المدن) :

متوسط الصف الأول : 150/6 = 25

متوسط الصف الثاني : 138/ 6 = 23

متوسط الصف الثالث : 126/ 6 = 21

متوسط الصف الرابع : 162/ 6 = 27

ب) المتوسط الكلي : $\frac{150+138+126+162}{24} = \frac{576}{24} = 24$

جـ) مجموع المربعات الكلي :

$$SST = (20^2 + 21^2 + 25^2 + 28^2 + 30^2 + 26^2 + 23^2 + \ldots + 27^2 + 24^2) - 6x4\,(24)^2$$

$$= 14160 - 13824$$

$$= 14160 - 6x4\,x24x24$$

$$SST = 336$$

د) مجموع المربعات للصفوف (المدن) :

$$SSR = 6(25^2 + 23^2 + 21^2 + 27^{2)} - 6x4x(24)^2$$

$$= 13944 - 13824$$

$$SSR = 120$$

هـ) مجموع مربعات الخطأ :

$$SSE = SST - SSR$$

$$= 336 - 120$$

$$SSE = 216$$

و) ثم نكون جدول تحليل التباين كما يلي : k=4, n=6

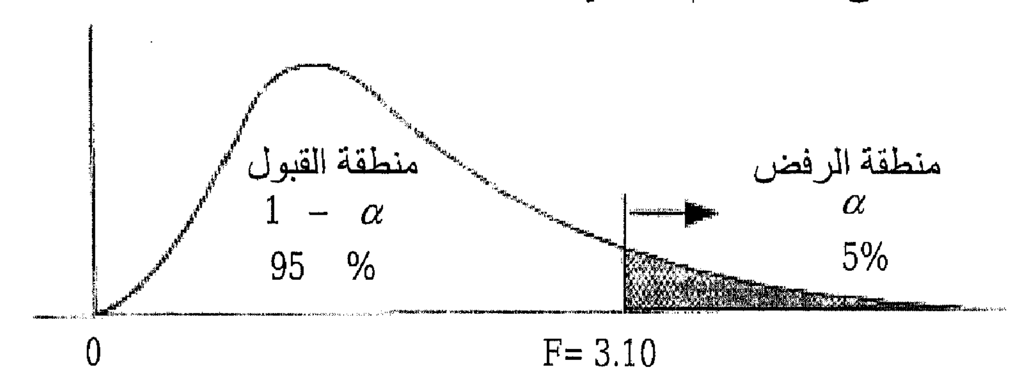
مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	الإحصائية F
بسبب المعاملات	SSR= 120	k-1 4-1 3	$S^2R = 120/3 =$ 40	$F = \frac{40}{10.8} = 3.7$
بسبب الخطأ	SSE= 216	K(n-1) 4(6-1) 20	$S^2E = 216/20 =$ 10.8	
الكلي	SST= 336	Nk-1 6*4-1 23		

أي أن قيمة الإحصائية (أو F المحسوبة) هي 3.7

$$F = \frac{40}{10.8} = 3.7$$

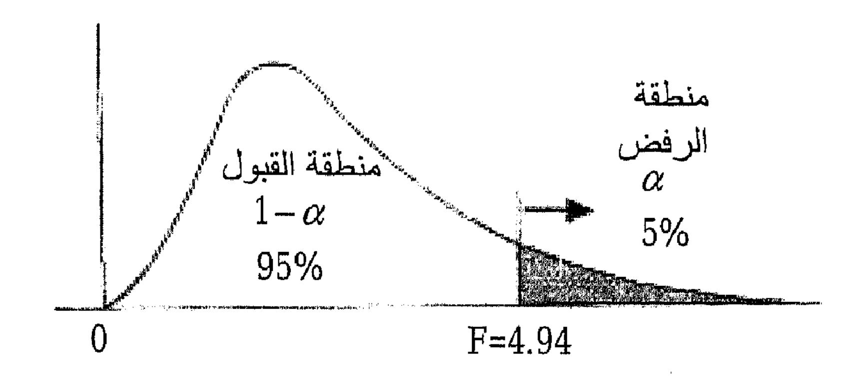
4- حـــدود منطقتي القبول والرفض : من جدول توزيع F وعند مستوى معنوية 5% وبدرجات حرية 3 للبسط، 20 للمقام نجد أن F الجدولية تساوي 3.10

ويمكن توضيح ذلك بالرسم كما يلي :



5- المقارنـــة والقرار: وحيث أن قيمة الإحصائية المسحوبة والتي تساوي 3.7 أكبر من القيمة الجدولية فإنها تقع في منطقة الرفض وبالتالي فإن القرار هو رفض الفرض العدمي بتساوي متوسطات أعمار الناخبين في المدن الأربع وذلك بمستوى معنوية 5%

* أما إذا استخدمنا مستوى معنوية 1% فإن قيمة F من الجدول تصبح 4.94 أي تصبح حدود منطقتي القبول والرفض كما يلي :



في هـــذه الحالة فإن قيمة الإحصائية والتي تساوي 3.7 تقع في منطقة القبول، وبالتالي فإن القرار يكون قبول الفرض العدمي يتساوى متوسطات أعمار الناخبين في المدن الأربع وذلك بمستوى معنوية 1%.

مـــثال9−2: في دراســـة لأثر طريقة التدريس في التحصيل في مادة الاحصاء لدى طلبة الجامعة. باستخدام ثلاثة طرق. كما هو موضح في الجدول أدناه:

الطريقة الأولى	الطريقة الثانية	الطريقة الثالثة
2	5	7
3	6	5
4	5	6
5	7	8
6	7	9

هل هناك فرق ذو دلالة احصائية في التحصيل يعزى لطريقة التدريس.

الحل:

متغيرات الدراسة:

المتغير المستقل: طريقة التدريس وتقسم إلى: الطريقة 1

الطريقة 2

الطريقة 3

المتغير التابع: التحصيل في مادة الاحصاء، ويقاس بالعلامة على الاختبار التحصيلي.

فرضيات الدراسة:

$H_0: \mu_1 = \mu_2 = \mu_3$ الفرضية الصفرية

$H_1: \mu_1 \neq \mu_2 \neq \mu_3$ الفرضية البديلة

الاقتـــراح: تجـــزئة الفرضية الصفرية الأم الى مجموعة من الفرضيات البسيطة واستخدام اختبار t في كل مرة، ويكون عدد الفرضيات الجزئية = $n(n-1)/2$

الاقتراح مرفوض للأسباب التالية:

1. لا يمكن من خلال اختبار t دراسة أثر التفاعل.
2. غير عملي في ظل زيادة عدد المتوسطات الحسابية.
3. تقل قيمة قوة الاختبار مع تثبيت قيمة α.
4. تزداد قيمة α (زيادة احتمال الوقوع في الخطأ) وهو شيء غير مرغوب فيه.

	الطريقة1		الطريقة2		الطريقة3			
	G1	G1^2	G2	G2^2	G3	G3^2	Total	Total^2
	2.00	4.00	5.00	25.00	7.00	49.00	2	4
	3.00	9.00	6.00	36.00	5.00	25.00	3	9
	4.00	16.00	5.00	25.00	6.00	36.00	4	16
	5.00	25.00	7.00	49.00	8.00	64.00	5	25
	6.00	36.00	7.00	49.00	9.00	81.00	6	36
							5	25
المجموع	20.00		30.00		35.00		6	36
المتوسط	4.00		6.00		7.00		5	25
Σx2	90.00		184.00		255.00		7	49
العدد	5		5		5		7	49
							7	49
							5	25
							6	36
							8	64
							9	81
المجموع							85	
المتوسط							5.67	
Σx2							529	
العدد							15	

لاختبار الفرضية الصفرية $H_0: \mu_1 = \mu_2 = \mu_3$

1- نحسب التباين بين المجموعات (Between Groups) وهو (MSB)

2- نحسب التباين ضمن المجموعات(Within Groups) وهي (MSW)

3- نحسب صيغة الاحصاءة المستخدمة لاختبار الفرضية هي:

$$F = \frac{MSB}{MSW} \quad .....F_{k-1,n-k}$$

درجة الحرية للبسط = (عدد المجموعات − 1)

درجة الحرية للمقام = (حجم العينة − عدد المجموعات)

4- القـــرار: إذا كانـــت قيمة F المحسوبة أكبر من القيمة الجدولية $F\alpha,\ k-1,\ k(n-1)$ نرفض الفرضية الصفرية التي تقول ان متوسطات المجموعات متساوية.

أي تباين في النهاية يصمم على شكل جدول كالتالي:

Source of Variance مصدر التباين	Sum of Sequares (SS) مجموع المربعات	Degrees of freedom (DF) درجات الحرية	Mean. Sequares (MS) متوسط المربعات	F ف
Between Groups بين المجموعات	SSB مجموع المربعات بين	C – 1 عدد الاعمدة-1	MSB=SSB/C-1 متوسط المربعات بين	MSB/MSW
Within Groups ضمن المجموعات	SSW مجموع المربعات ضمن	N – C حجم العينة-عدد الاعمدة	MSW=SSW/N-C متوسط المربعات ضمن	
Total الكلي	SST مجمـــوع المربعات الكلي	N – 1 حجم العينة-1		

التباين داخل المجموعات هو تباين عشوائي (ناتج عن اختلاف الأفراد).

التـــباين بين المجموعات هو تباين عشوائي (ناتج عن اختلاف الأفراد)، وتباين ناتج عن اختلاف طريقة التدريس.

عندما يكون التباين لطريقة التدريس = 0، فإن التباين الداخلي=التباين الخارجي، وهذا يحدث في حالة أن الفرضية الصفرية صحيحة.

حجم العينة N ، عدد المجموعات في العينة C

$MSW = SSW / N\text{-}C$
$MSB = SSW / C\text{-}1$
$SST = SSB + SSW$

احسب الكميات:

طريقة 1:

$A = \Sigma x^2$
$B = (\Sigma x)^2/N$
$C = (\Sigma x)^2/n_1 + (\Sigma x)^2/n_2 + \ldots.$
$SST = A - B$
$SSB = C - B$
$SSW = A - C$

طريقة 2:

$SST = \Sigma(X\text{-}X')^2$ مجموع مربعات الانحراف عن الوسط الكلي

$SSB = \Sigma n_i(X'_i\text{-}X')^2$

$SSW = \Sigma(X\text{-}X'_i)^2$ لكل المجموعات

طريقة 3: اذا علمت المتوسطات الحسابية والانحرافات المعيارية.

$SSB = n_1 (X'_1 - X')^2 + n_2 (X'_2 - X')^2 + \ldots$
$SSW = (n_1\text{-}1)S_1^2 + (n_2\text{-}1)S_2^2 + \ldots..$

حل المثال 9-2

$A = \Sigma x^2 = 529$
$B = (\Sigma x)^2/N = (85)^2/15 = 481.67$
$C = (\Sigma x)^2/n_1 + (\Sigma x)^2/n_2 + \ldots. = (20)^2/5 + (30)^2/5 + (35)^2/5 = 505$
$SST = A - B = 529 - 481.67 = 47.33$
$SSB = C - B = 505 - 481.67 = 23.33$
$SSW = A - C = 529 - 505 = 24$

Source of Variance مصدر التباين	Sum of Sequares (SS) مجمـــــــوع المربعات	Degrees of freedom (DF) درجات الحرية	Mean. Sequares (MS) متوسط المربعات	F ف
Between Groups بين المجموعات	SSB 23.33	C – 1 3-1 = 2	MSB=SSB/C-1 23.33/2= 11.665	MSB/MSW 11.665/2 = 5.8325
Within Groups ضمن المجموعات	SSW 24	N – C 15-3 = 12	MSW=SSW/N-C 24/12 = 2	
Total الكلي	SST 47.33	N – 1 15-1 = 14		

القرار:

القيمة المحسوبة لـــ F = 5.8325

القيمة الحرجة (الجدولية) F2,12,0.05 = 3.89

بمـــا أن القيمة المحسوبة > أكبر من القيمة الحرجة نرفض الفرضية الصفرية، ونستنتج أنه يوجد اختلاف في التحصيل يعزى الى طريقة التدريس.

ثم نستخرج الجدول العام للنسبة الفائية

مصدر التباين	مجموع المربعات		درجة الحرية		التباين	النسبة الفائية
بين المجموعات	441.17	÷	3	=	147.06	
					÷	14.29
داخل المجموعات	164.63	÷	16	=	10.29	
المجموع	605.8		19			

ف=14.29، دح=16/3، دالة عند 0.01 (الرجوع الى جدول النسبة الفائية لقراءته)

كيفية قراءة جدول النسبة الفائية حسب النسبة الفائية التي حصلنا عليها وحسب درجات الحرية التي لدينا.

TABLE XI

TABLE DU F DE SNEDECOR

Les deux valeurs de F correspondent respectivement aux seuils de P = .05 et P = .01.

ن1

درجة الحرية ن1 و ن2

ن2

حدود الدلالة عند 0.05 أي السطر الأول

حدود الدلالة عند 0.01 أي السطر الثاني

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
2	18,51 98,49	19,00 99,00	19,16 99,17	19,25 99,25	19,30 99,30	19,33 99,33	19,37 99,36	19,41 99,42	19,45 99,46	19,50 99,50
3	10,13 34,12	9,55 30,82	9,28 29,46	9,12 28,71	9,01 28,24	8,94 27,91	8,84 27,49	8,74 27,05	8,64 26,60	8,53 26,12
4	7,71 21,20	6,94 18,00	6,59 16,69	6,39 15,98	6,26 15,52	6,16 15,21	6,04 14,80	5,91 14,37	5,77 13,93	5,63 13,46
5	6,61 16,26	5,79 13,27	5,41 12,06	5,19 11,39	5,05 10,97	4,95 10,67	4,82 10,27	4,68 9,89	4,53 9,47	4,36 9,02
6	5,99 13,74	5,14 10,92	4,76 9,78	4,53 9,15	4,39 8,75	4,28 8,47	4,15 8,10	4,00 7,72	3,84 7,31	3,67 6,88
7	5,59 12,25	4,74 9,55	4,35 8,45	4,12 7,85	3,97 7,46	3,87 7,19	3,73 6,84	3,57 6,47	3,41 6,07	3,23 5,65
8	5,32 11,26	4,46 8,65	4,07 7,59	3,84 7,01	3,69 6,63	3,58 6,37	3,44 6,03	3,28 5,67	3,12 5,28	2,93 4,86
9	5,12 10,56	4,26 8,02	3,86 6,99	3,63 6,42	3,48 6,06	3,37 5,80	3,23 5,47	3,07 5,11	2,90 4,73	2,71 4,31
10	4,96 10,04	4,10 7,56	3,71 6,55	3,48 5,99	3,33 5,64	3,22 5,39	3,07 5,06	2,91 4,71	2,74 4,33	2,54 3,91
12	4,75 9,33	3,88 6,93	3,49 5,95	3,26 5,41	3,11 5,06	3,00 4,82	2,85 4,50	2,69 4,16	2,50 3,78	2,30 3,36
14	4,60 8,86	3,74 6,51	3,34 5,56	3,11 5,03	2,96 4,69	2,85 4,46	2,70 4,14	2,53 3,80	2,35 3,43	2,13 3,00
16	4,49 8,53	3,63 6,23	3,24 5,29	3,01 4,77	2,85 4,44	2,74 4,20	2,59 3,89	2,42 3,55	2,24 3,18	2,01 2,75
18	4,41 8,28	3,55 6,01	3,16 5,09	2,93 4,58	2,77 4,25	2,66 4,01	2,51 3,71	2,34 3,37	2,15 3,00	1,92 2,57
20	4,35 8,10	3,49 5,85	3,10 4,94	2,87 4,43	2,71 4,10	2,60 3,87	2,45 3,56	2,28 3,23	2,08 2,86	1,84 2,42
25	4,24 7,77	3,38 5,57	2,99 4,68	2,76 4,18	2,60 3,86	2,49 3,63	2,34 3,32	2,16 2,99	1,96 2,62	1,71 2,17
30	4,17 7,56	3,32 5,39	2,92 4,51	2,69 4,02	2,53 3,70	2,42 3,47	2,27 3,17	2,09 2,84	1,89 2,47	1,62 2,01
40	4,08 7,31	3,23 5,18	2,84 4,31	2,61 3,83	2,45 3,51	2,34 3,29	2,18 2,99	2,00 2,66	1,79 2,29	1,51 1,81
60	4,00 7,08	3,15 4,98	2,76 4,13	2,52 3,65	2,37 3,34	2,25 3,12	2,10 2,82	1,92 2,50	1,70 2,12	1,39 1,60
80	3,96 6,96	3,11 4,88	2,72 4,04	2,48 3,56	2,33 3,25	2,21 3,04	2,05 2,74	1,88 2,41	1,65 2,03	1,32 1,49

3-9 تحليل التباين الثنائي Tow-Way Analysis of Variance

تحليل التباين الثنائي Tow-Way Analysis of Variance:

يستخدم لدراسة أثر متغيرين عاملين يقسم كل منهما أفراد العينة إلى مستويين أو أكثر على متغير كمي (المتغير التابع).

يستخدم لمقارنة عدة متوسطات حسابية بهدف دراسة أثر متغيرين مستقلين على متغير تابع واحد.

أي يستخدم تحليل التباين الثنائي لدراسة الأثر لعاملين هما العامل a والعامل b على متغير تابع واحد، مثل دراسة أثر طريقة التدريس (العامل a)، والجنس (العامل b)، على التحصيل (المتغير التابع).

ويقصد بأثر التفاعل a X b على المتغير التابع هو اختلاف تأثير a على المتغير التابع من مستوى الى آخر من مستويات b ، أي يختلف تأثير طريقة التدريس على التحصيل لدى الذكور عنه لدى الإناث، وفحص التفاعل مهم جداً حيث يعني وجوده التخصيص في النتائج، ويعني غيابه تعميم النتيجة.

إن تحليل التباين الأحادي يستخدم لدراسة أثر عامل واحد (المتغير العاملي) على متغير ما. ولكن ماذا لو اردنا دراسة أثر عاملين أو اكثر على متغير ما؟ في هذه الحالة يمكننا استخدام تحليل التباين الثنائي أو الثلاثي.

كثيرا ما يواجه الباحثون مشاكل تتطلب دراسة نوعين من العوامل أو المعالجات وفي هذه الحالة يتم إجراء تحليل التباين بوجود معيارين للتصنيف فمثلا يكون المعيار الأول (طلاب في مدرسة) والمعيار الثاني (طرق تدريس مختلفة) فإذا أردنا معرفة ما إذا كان الاختلاف بين طلاب المدرسة عائد على اختلاف طريقة التدريس المتبعه معهم، وقد نرغب في معرفة التأثير المشترك لكلا العاملين (أي تأثير التفاعل). وفي الدراسات التطبيقية التجريبية يمكن مثلاً دراسة تأثير التربة ونوعية السماد المستخدم في انتاج القمح، أو دراسة تأثير جودة مواد البناء ونوعية المهندسين لعمل البيوت السكنية، أو دراسة تأثير مناطق بيع البضائع ومصاريف الدعاية على كمية المبيعات.

فتحليـــيل التـــباين الثنائي **Two Way ANOVA** يمكن استخدامه لدراسة اثر متغيرين عـــاملين يقسم كل منهما مفردات العينة الى مستويين (مجموعتين) او أكثر على متغير كمي ما (المتغير التابع). ويستخدم في هذه الحالة جداول مزدوجة مكونة من صفوف وأعمدة وتضم كل خلية من خلاياهما مشاهدتين أو أكثر (مفردات تنتمي إلى صف معين وعمود بنفس الوقت).

- إن تحليل التباين باتجاهين يستخدم لمعرفة إن كانت هناك فروق معنوية بين المتوسطات (متوســـطات معالجات الصفوف، بين متوسطات معالجات الأعمدة، وكذلك بين متوسطات المشاهدات الموجودة في كل خلية حيث تستخدم متوسطات الخلايا لاختبار التفاعل).

من خلال تحليل التباين الثنائي يمكن اختبار ثلاث فرضيات كما يلي:

1. الأثر الرئيسي للمتغير العاملي الأول على المتغير التابع. $H0: \mu a1= \mu a2= \mu a3$
2. الأثر الرئيسي للمتغير العاملي الثاني على المتغير التابع. $H0: \mu b1= \mu b2= \mu b3$
3. أثر التفاعل بين المتغيرين العامليين على المتغير التابع. a*b

$H0: \Sigma (\alpha B)^2 = 0$

شروط تحقيق التباين الثنائي:

- يجب أن يكون توزيع المتغير التابع طبيعيا لكل مجتمع من المجتمعات في تصميم التجربة، أي ان كل مجتمع ممثل بكل خلية من خلايا تصميم التجربة، فإذا كان على سبيل المثال ثـــلاث مستويات لكل متغير عاملي فيكون هناك 9 خلايا. وان لم يتحقق هذا الشرط فانـــه يمكن الاستغناء عنه بزيادة حجم العينة بحيث تزيد على 15 مفردة لكل مجموعة (خلـــية) ، وفي هذه الحالة قد تكون نتيجة تحليل التباين دقيقة الى حد ما حتى لو كان توزيع المتغير التابع ليس طبيعيا.
- يجب ان يكون تباين المتغير التابع متساويا لكل مجتمع من مجتمعات المعرفة في كل خلية من خلايا تصميم التجربة، وإذا لم يتحقق هذا الشرط فإن نتيجة تحليل التباين لن تكون دقيقة. أما المقارنات البعدية الخاصة بالأثر الرئيسي فمن الممكن استخدام بعض الطرائق التي لا تشترط تساوي التباين .
- يجب أن تكون العينات مختارة بطريقة عشوائية من كل مجتمع من المجتمعات. ويجب أن تكون قيم المتغير التابع مستقلة عن بعضها بعضا لكل مفردة من مفردات العينات.

هناك نماذج متعددة لتحليل التباين الثنائي ترتبط بالتصميم التجريبي الذي يعتمده الباحث منها:

1. تحليل التباين بآثار ثابتة Fixed Effects
2. تحليل التباين بآثار عشوائية Random Effects
3. تحليل التباين بآثار خليط Mixed Effects

تحليل التباين بآثار ثابتة Fixed Effects

مثال:

المتغيرات المستقلة: 1- الجنس: وله مستويان ذكر وانثى.

2- طرق تعليم القراءة: ولها ثلاثة مستويات.

المتغير التابع : القدرة على الاستيعاب.

*** تحليل التباين الثنائي Two Way Analysis of Variance**

تحليل التباين الثنائي Two Way ANOVA يمكن استخدامه لدراسة اثر متغيرين عاملين يقسم كل منهما مفردات العينة الى مستويين (مجموعتين) او اكثر على متغير كمي ما (المتغير التابع).

مثال:

باستخدام ملف Employee data " الموجود مع برمجية SPSS اختبر الفرضية التالية:

" لا يؤثر الجنس ونوع العمل في تحديد الراتب للموظفين بمستوى معنوية 0.05 "

الحل:

المتغيرات المستقلة: إن هناك عاملان يؤثران على تحديد الراتب هما:

1- الجنس. 2- نوع العمل.

المتغير التابع: الراتب.

لذلك يمكن تقسيم هذه الفرضية الى ثلاث فرضيات جزئية وهي:

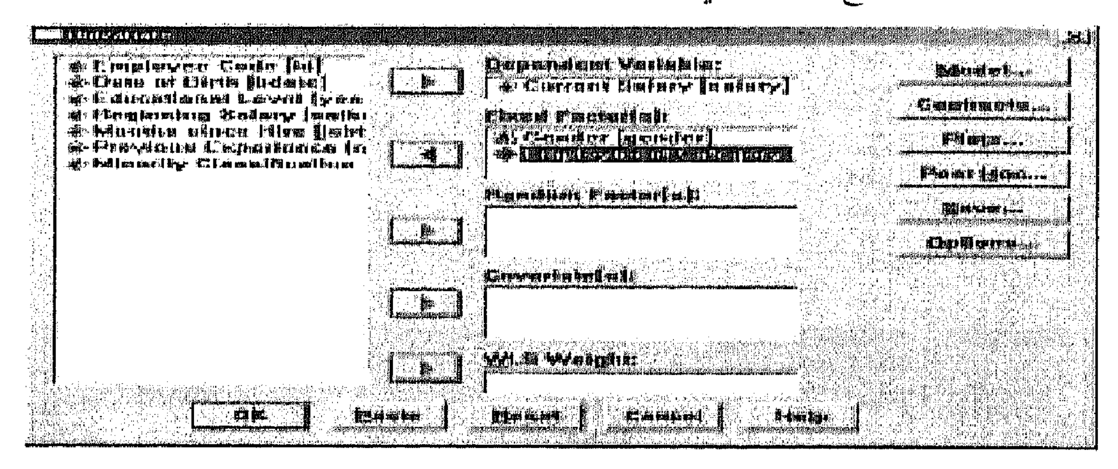
الفرضية الاولى " *لا تاثير للجنس في تحديد الراتب* "

الفرضية الثانية "*لا تأثير لنوع العمل على تحديد الراتب* "

الفرضية الثالثة" *لا يوجد تفاعل بين متغير الجنس ومتغير نوع العمل* "

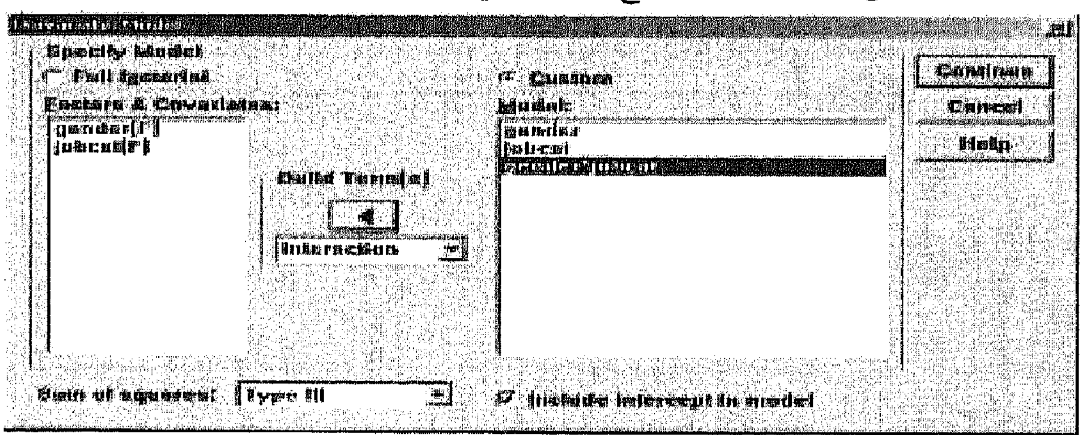
ولفحص الفرضيات نستخدم تحليل التباين الثنائي كما يلي:

1. مـــن القائمة Analyze اختر General Linear Model ومن القائمة الفرعية اختر Univariate يظهر مربع الحوار التالي:



2. انقـــل المـــتغير Salary الى المـــستطيل أسفل Dependent Variable والمتغيران Gender وJobcat إلى المستطيل أسفل Fixed Factor(s) .

3. اضغط على Model يظهر مربع الحوار التالي:



4. اختـــر Custom ثم Main effects مـــن القائمة أسفل Build Term (s) وانقل المتغيرين Gender و Jobcat الى المستطيل أسفل Model ، ثم اختر Interaction من القائمة Build Term وانقـــل المـــتغيرين , Gender و Jobcat معا إلى المستطيل أسفل Model ، اضغط Continue سنعود الى المربع الاصلي.

5. اضـــغط Options سيظهر مربع الحوار التالي اختر منه Descriptive atatistics والخيار Homogeneity tests ثم اضغط Continue لنعود لمربع الحوار الأصلي.

Descriptive Statistics

Dependent Variable: Current Salary

Gender	Employment Category	Mean	Std. Deviation	N
Female	Clerical	$25,003.69	$5,812.838	206
	Manager	$47,213.50	$8,501.253	10
	Total	$26,031.92	$7,558.021	216
Male	Clerical	$31,558.15	$7,997.978	157
	Custodial	$30,938.89	$2,114.616	27
	Manager	$66,243.24	$18,051.570	74
	Total	$41,441.78	$19,499.214	258
Total	Clerical	$27,838.54	$7,567.995	363
	Custodial	$30,938.89	$2,114.616	27
	Manager	$63,977.80	$18,244.776	84
	Total	$34,419.57	$17,075.661	474

6. اضغط على Post Hoc ليظهر مربع الحوار التالي:

7. اختـــر اختبار شفيه Scheffe للمقارنات البعدية من قائمة الاختبارات البعدية التي تشترط تماثل تباينات الفئات Equal Variance Assumed .

8. اختر اختبار دونت س C Dunnett,s من قائمة الاختبارات البعدية التي لا تشترط تماثل تباينات الفئات Equal Variance Not Assumed

9. انقل المتغير Jobcat فقط الى المستطيل اسفل Post Hoc Tests For لأنه يتكون من ثلاث مستويات أما متغير Gender فلا ننقله لأنه يتكون من مستويين فقط.

10. اضغط Continue سنعود لمربع الحوار الأصلي.

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Gender	f	Female	216
	m	Male	258
Employment Category	1	Clerical	363
	2	Custodial	27
	3	Manager	84

الجدول التالي يبين توزيع العينة حسب مستويات كل من المتغيرات العاملية.

جدول اختبار تجانس التباينtest of Homogeneity of Variances ، ويبين أن قيمة Sig. =0.0 يعني أن تباين المجموعات غير متساو لأنها اكبر من 0.05.

الجدول التالي يبين تحليل التباين الثنائي حسب فئات المتغير Gender ، ويظهر أن Sig. = 0.0 وهـــي اقل من 0.05 أي أن الجنس يؤثر في تحديد الراتب. كذلك الجدول التالي يبين تحليل التباين الثنائي حسب فئات المتغير Jobcat ، ويظهر أن Sig=0.0 وهي اقل من 0.05 أي أن نوع العمل يؤثر في تحديد الراتب. كما يظهر أن هناك تفاعل بين الجنس ونوع الوظيفة لان قيمة Sig=0.0 وهي اقل من 0.05

Levene's Test of Equality of Error Variances [a]

Dependent Variable: Current Salary

F	df1	df2	Sig.
33.383	4	469	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+GENDER+JOBCAT+GENDER * JOBCAT

Multiple Comparisons

Dependent Variable: Current Salary

	(I) Employment Category	(J) Employment Category	Mean Difference (I-J)	Std. Error	Sig.
Scheffe	Clerical	Custodial	-$3,100.35	$1,875.539	.256
		Manager	-$36,139.26*	$1,138.387	.000
	Custodial	Clerical	$3,100.35	$1,875.539	.256
		Manager	-$33,038.91*	$2,080.027	.000
	Manager	Clerical	$36,139.26*	$1,138.387	.000
		Custodial	$33,038.91*	$2,080.027	.000
Dunnett C	Clerical	Custodial	-$3,100.35*	$568.679	
		Manager	-$36,139.26*	$2,029.912	
	Custodial	Clerical	$3,100.35*	$568.679	
		Manager	-$33,038.91*	$2,031.840	
	Manager	Clerical	$36,139.26*	$2,029.912	
		Custodial	$33,038.91*	$2,031.840	

Based on observed means.

*. The mean difference is significant at the .05 level.

الجـــدول التالي يبين أن متوسطات Jobcat مختلفة ويبين أن متوسطات الكتاب المدراء وكذلك متوسطات الحراس والمدراء مختلفة بينما لا يوجد خلاف بين متوسطات رواتب الحراس والكتاب له دلالة إحصائية تذكر

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.646E+10[a]	4	2.411E+10	272.780	.000
Intercept	1.773E+11	1	1.773E+11	2005.313	.000
GENDER	5247440732	1	5247440732	59.359	.000
JOBCAT	3.232E+10	2	1.616E+10	182.782	.000
GENDER * JOBCAT	1247682867	1	1247682867	14.114	.000
Error	4.146E+10	469	88401147.44		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .699 (Adjusted R Squared = .697)

* تحليل التباين الثلاثي Three Way ANOVA

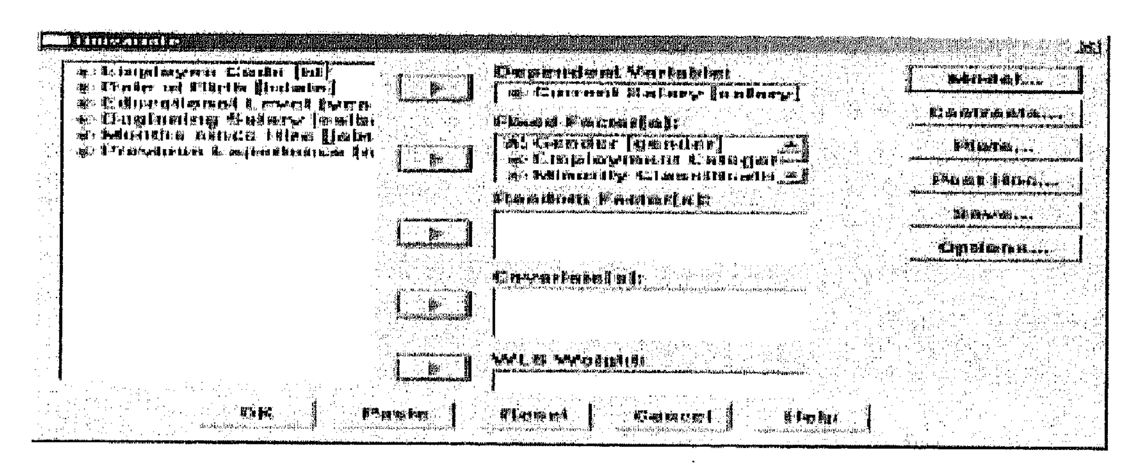
اذا كـــان لديـــنا ثلاث متغيرات عاملية واردنا فحص اثر هذه العوامل على متغير تابع نستخدم تحليل التباين الثلاثي ونتبع نفس خطوات تحليل التباين الثنائي ولناخذ المثال التالي:

مثال : استخدم ملف Employee data لفحص الفرضية التالية :

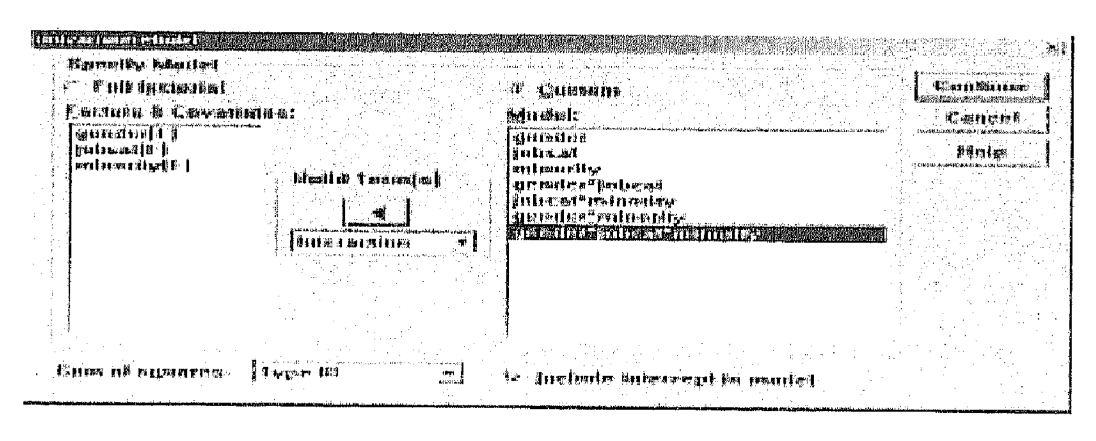
" لا يوجد فرق في متوسطات رواتب الموظفين تحت تاثير الجنس و نوع الوظيفة والاقلية بمستوى دلالة 0.05 "

ولفحص هذه الفرضية نتبع الخطوات التالية:

1. مـــن القائمة Analyze اختر General Linear Model ومن القائمة الفرعية اختر Univariate يظهر مربع الحوار التالي:



2. انقـــل المـــتغير Salary الى المـــستطيل اسفل Dependent Variable والمتغيرات Gender و id وJobcat الى المستطيل اسفل Fixed Factor(s) .
3. اضغط على Model يظهر مربع الحوار التالي:



4. اختــر Custom ثم Main effects مــن القائمة أسفل Build Term (s) وانقل المتغيرين Gender و Jobcat و id الى المستطيل أسفل Model ، ثم اختر Interaction من القائمــة Build Term وانقل المتغيرات معا مثنى مثني ثم جميعهم إلى المستطيل اسفل Model كما بالشكل أعلاه، اضغط Continue سنعود إلى المربع الاصلي.

5. ضــغط Options سيظهر مربع الحوار التالي اختر من Descriptive atatistics والخيار Homogeneity tests وانقل المتغيرات الثلاثة الى المستطيل أسفل Display Means for ثم اضغط Continue لنعود لمربع الحوار الأصلي.

6. اضغط على Post Hoc ليظهر مربع الحوار التالي:

7. اختــر اختبار شفيه Scheffe للمقارنات البعدية من قائمة الاختبارات البعدية التي تشترط تماثل تباينات الفئات Equal Variance Assumed .

8. اختـــر اختبار دونت س Dunnett,s C من قائمة الاختبارات البعدية التي لا تشترط تماثل تباينات الفئات Equal Variance Not Assumed

9. انقل المتغير Jobcat فقط إلى المستطيل اسفل Post Hoc Tests For لأنه يتكون من ثلاث مستويات أما المتغيران الاخران فلا ننقلهما لأنهما يتكونان من مستويين فقط.

10. اضغط Continue سنعود للمربع الحوار الأصلي. اضغط Ok تائج التالية:

على الدارس تفسير النتائج

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Minority Classification	0	No	370
	1	Yes	104
Gender	f	Female	216
	m	Male	258
Employment Category	1	Clerical	363
	2	Custodial	27
	3	Manager	84

Descriptive Statistics

Dependent Variable: Current Salary

Minority Classificat	Gender	Employment Categ	Mean	Std. Deviation	N
No	Female	Clerical	25,471.45	$6,092.372	166
		Manager	47,213.50	$8,501.253	10
		Total	26,706.79	$8,011.894	176
	Male	Clerical	32,671.64	$8,578.999	110
		Custodial	31,178.57	$1,658.743	14
		Manager	65,683.57	$18,029.451	70
		Total	44,475.41	$20,330.662	194
	Total	Clerical	28,341.09	$7,994.659	276
		Custodial	31,178.57	$1,658.743	14
		Manager	63,374.81	$18,164.043	80
		Total	36,023.31	$18,044.096	370
Yes	Female	Clerical	23,062.50	$3,972.369	40
		Total	23,062.50	$3,972.369	40
	Male	Clerical	28,952.13	$5,712.419	47
		Custodial	30,680.77	$2,562.920	13
		Manager	76,037.50	$17,821.961	4
		Total	32,246.09	$13,059.881	64
	Total	Clerical	26,244.25	$5,772.874	87
		Custodial	30,680.77	$2,562.920	13
		Manager	76,037.50	$17,821.961	4
		Total	28,713.94	$11,421.638	104
Total	Female	Clerical	25,003.69	$5,812.838	206
		Manager	47,213.50	$8,501.253	10
		Total	26,031.92	$7,558.021	216
	Male	Clerical	31,558.15	$7,997.978	157
		Custodial	30,938.89	$2,114.616	27
		Manager	66,243.24	$18,051.570	74
		Total	41,441.78	$19,499.214	258
	Total	Clerical	27,838.54	$7,567.995	363
		Custodial	30,938.89	$2,114.616	27
		Manager	63,977.80	$18,244.776	84
		Total	34,419.57	$17,075.661	474

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Current Salary

F	df1	df2	Sig.
17.696	8	465	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+MINORITY+GENDER+JOBCAT+MINORITY * GENDER+GENDER * JOBCAT+MINORITY * JOBCAT+MINORITY * GENDER * JOBCAT

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.751E+10[a]	8	1.219E+10	140.251	.000
Intercept	1.444E+11	1	1.444E+11	1661.526	.000
MINORITY	61989119.7	1	61989119.66	.713	.399
GENDER	4756876310	1	4756876310	54.737	.000
JOBCAT	2.006E+10	2	1.003E+10	115.420	.000
MINORITY * GEND	27977363.9	1	27977363.93	.322	.571
GENDER * JOBCAT	981526336	1	981526335.9	11.294	.001
MINORITY * JOBCA	690053398	2	345026699.0	3.970	.020
MINORITY * GEND * JOBCAT	.000	0	.	.	.
Error	4.041E+10	465	86903667.84		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .707 (Adjusted R Squared = .702)

Estimated Marginal Means

1. Minority Classification

Dependent Variable: Current Salary

Minority Classification	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
No	40443.745[a]	835.531	38801.861	42085.629
Yes	39683.224[a]	1423.737	36885.469	42480.979

a. Based on modified population marginal mean.

2. Employment Category

Dependent Variable: Current Salary

Employment Category	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Clerical	27539.427	577.449	26404.695	28674.160
Custodial	30929.670[a]	1795.293	27401.779	34457.562
Manager	62978.190[a]	1875.508	59292.670	66663.711

a. Based on modified population marginal mean.

Post Hoc Tests
Employment Category

Multiple Comparisons

Dependent Variable: Current Salary

	(I) Employment Category	(J) Employment Category	Mean Difference (I-J)	Std. Error	Sig.
Scheffe	Clerical	Custodial	-$3,100.35	$1,859.586	.250
		Manager	-$36,139.26*	$1,128.703	.000
	Custodial	Clerical	$3,100.35	$1,859.586	.250
		Manager	-$33,038.91*	$2,062.334	.000
	Manager	Clerical	$36,139.26*	$1,128.703	.000
		Custodial	$33,038.91*	$2,062.334	.000
Dunnett C	Clerical	Custodial	-$3,100.35*	$568.679	
		Manager	-$36,139.26*	$2,029.912	
	Custodial	Clerical	$3,100.35*	$568.679	
		Manager	-$33,038.91*	$2,031.840	
	Manager	Clerical	$36,139.26*	$2,029.912	
		Custodial	$33,038.91*	$2,031.840	

Based on observed means.

*. The mean difference is significant at the .05 level.

4-9 تحليل التغاير (ANCOVA) nalysis of Covariance

يعتبر تحليل التغاير (ANCOVA) من الأساليب المستخدمة مع التصميم شبه التجريبي والفرق بينهما انه في حال التصميم التجريبي يتم ضبط المتغيرات المتحكمة في المتغير التابع أما في التصميم شبه التجريبي فإن بعض المتغيرات المؤثرة في النتائج لم يتم إجراء تكافؤ بين المجموعات التجريبية والضابطة إنما تم قياس تلك المتغيرات فقط وبالتالي فإن استخدام الاختبار البديل وهو تحليل التباين لن يعطي نتائج حقيقية ، لذلك نستخدم أختبار تحليل التغاير.

مثال: تم اختيار 9 مفحوصين عشوائياً و توزيعهم علي 3 مجموعات درست كل مجموعة منهم بإحدي طرق التدريس الرياضيات وتم قياس الاستعداد الرياضي والتحصيل حيث تم قياس الاستعداد الرياضي أو القدرة الرياضية قبل التجرية وقياس التحصيل بعد التجريب.

مجموعة 1		مجموعة 2		مجموعة 3	
استعداد	تحصيل	استعداد	تحصيل	استعداد	تحصيل
Apt	Ach	Apt	Ach	Apt	Ach
8	6	7	8	9	12
4	3	5	7	10	12
6	4	8	10	7	10

ونلاحظ أن البيانات يجب إدخالها بالترتيب التالي ببرنامج SPSS

Methodالطريقة	Aptاستعداد	Achتحصيل
1	8	6
1	4	3
1	6	4
2	7	8
2	5	7
2	8	10
3	9	12
3	10	12
3	7	10

الآن لدينا ثلاثة متغيرات وهي:

- المتغير المستقل : طريقة بثلاث مستويات أو ثلاثة مجموعات مستقلة .
- متغير التغاير : اختبار القدرة الرياضية .
- المتغير التابع : اختبار التحصيل في الرياضيات

والواقع أن اختبار انكوفا يحاول حذف أثر المتغير المصاحب من المتغير التابع – التحصيل ومن ثم نستخدم اختبار – ف F علي المتوسطات المعدلة للاختبار التحصيلي لتبين الفروق بينها، وفي حال وجود أكثر من طريقتين أو عينتين يجب استخدام اختبارات المقارنة البعدية

وبرنامج SPSS لديه LSD و ت- بونفروني وSidak المعدل لتحسس موضع الفروق ، ولكن يفضل اختبار بونفروني و Sidak لأنهما يعتمدان علي المتوسطات المعدلة.

شكل البيانات بمحرر SPSS

	method	apt	ach
1	1.00	8.00	6.00
2	1.00	4.00	3.00
3	1.00	6.00	4.00
4	2.00	7.00	8.00
5	2.00	5.00	7.00
6	2.00	8.00	10.00
7	3.00	9.00	12.00
8	3.00	10.00	12.00
9	3.00	7.00	10.00

التحقق من فرضية تساوي ميل الانحدار :

حتي نطبق تحليل التغير يجب أن تكون العلاقة بين متغير التغاير والمتغير التابع خطية وعدم وجود تفاعل بين متغير التغاير و المعالجات

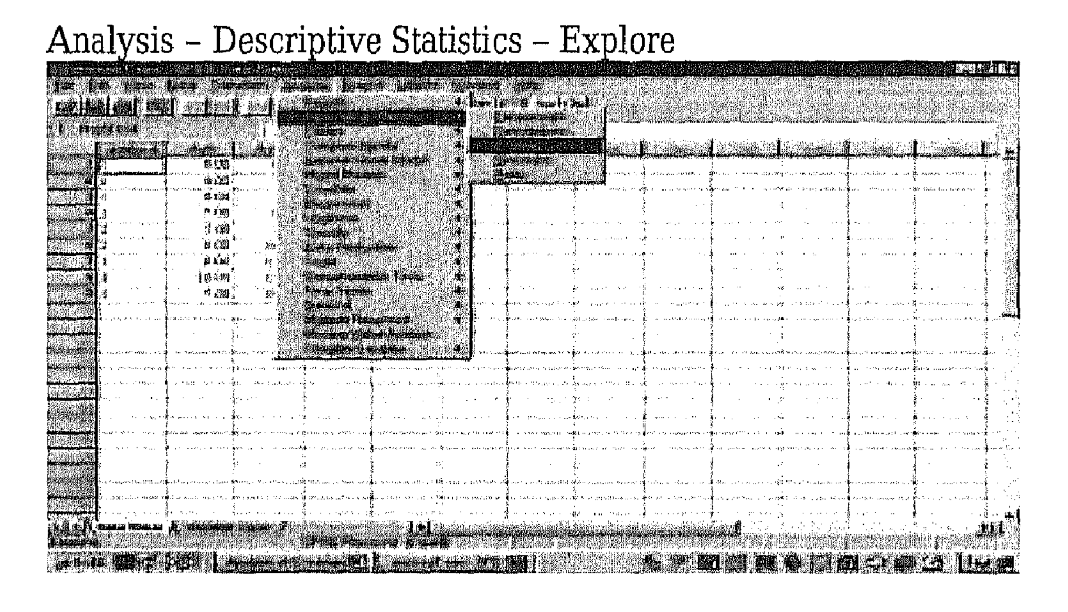
هل توجد علاقة خطية قوية بين متغير التغاير و المتغير التابع ؟هل ميل الخطوط لكل من متغير التغاير – الاستعداد – ومتغير التحصيل في المجموعات الثلاث متساوي ؟ (هل التفاعل بين الطرق الثلاث و الاستعداد دال إحصائيا؟

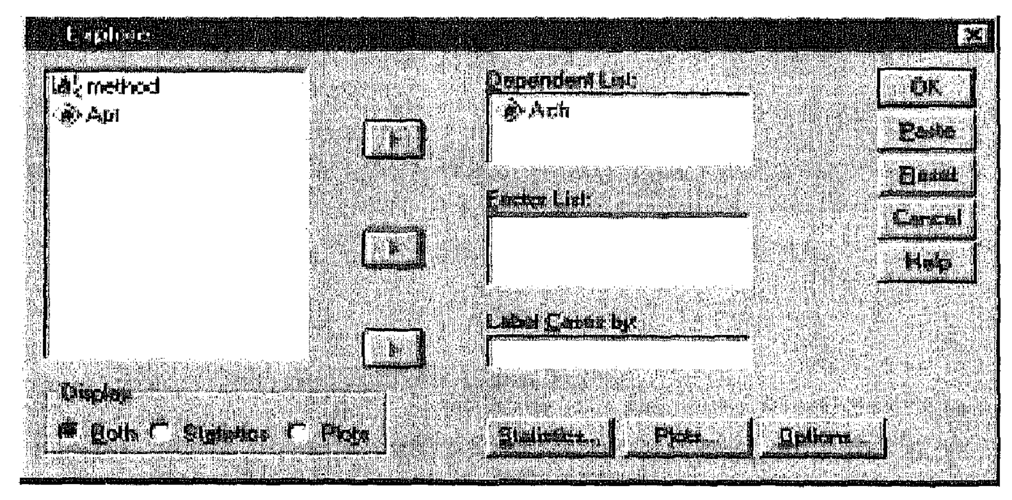
§ **الخطية** : نظراً لأن اختبار أنكوفا عبارة عن نموذج خطي عام مع انحدار متعدد ، فيفترض أن متغير التغاير يرتبط خطياً مع المتغير التابع ، وفي مثالنا يجب أن يكون الارتباط خطي بين كل من الاستعداد الرياضي والتحصيل في الرياضيات ، وفي برنامج SPSS يجب أن نرسم أو نحدد الانحدار الخطي لكل المجموعات بين المتغيرين ثم نحسب مربع معامل الارتباط ونقارن بين الميل في كل مجموعة ..

§ **تجانس التباين** : يجب أن يكون التباين متجانساً لكل المجموعات بين المتغيرين وهو نفس شرط اختبار أنوفا مع ملاحظة شرط الاعتدالية أيضاً ، وشرط الاعتدالية شرط رياضي مهم ، وقد اقترح بعض الباحثين أن هذا الشرط يمكن التغاضي عنه في حال العينات التي تزيد

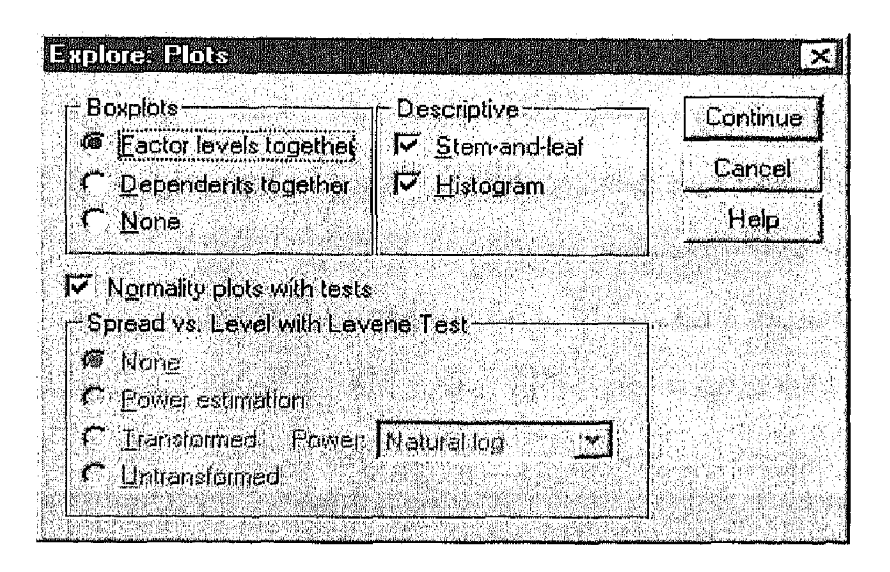
عـــن 30 وفي مثالنا العينة 3 ولكن فؤاد أبوحطب نقلا عن باحثين أخرين ذكر بتقليص العدد لـــ 15 علي أساس أن الفروق تكون ضئيلة ، ولكن برنامج مثل sigmaStat يقوم بإجراء هذه العملـــية في كـــل مرة وللأسف برنامج SPSS يترك ذلك لخبرة مشغل البرنامج الذي عليه ان يـــتحقق مـــن شرطي التجانس وشرط الاعتدالية بنفسه باستخدام اختبار ليفين لتجانس التباين واختبار لليفورأو اختبار Shapiro-Wilk أو اختبار كلمجروف – سميرانوف Kolmogorov-Smirnov ونستطيع إجراهما من خلال برنامج SPSS وفق الخطوات التالية :

Analysis – Descriptive Statistics – Explore





انقر الزر Plots تحصل علي النافذة التالية:



علـــم خانـــة Normality plots with tests وانقر الزر Continue تظهر النافذة التي تسبقها ثم انقر موافق .

Tests of Normality

Shapiro-Wilk			Kolmogorov-Smirnov(a)			
Sig.	Df	Statistic	Sig.	df	Statistic	
.537	9	.936	.200(*)	9	.174	Ach

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

ويخـــشي بعض الباحثين عدم تحقق شرط الاعتدالية والحل ببساطة هو اللجوء للإحصاء الابارامترك.

§ **تجانس الانحدار** : نفترض أنكوفا تجانس الانحدار وبعض البرامج تحسبه بشكل مستقل ولكـــن يـــستدل عليه في برنامج SPSS بحساب دلالة التفاعل بين المتغيرين من خلال حساب جدول أنكوفا مع حساب التفاعل.

حساب تجانس الانحدار مع باستخدام برنامج SPSS

1. اختر Analysis - General Linear Model - Univariate.

§ اختر المتغير ach وضعه في مربع المتغير التابع Independent variable .

▪ اختر method وضعها في مربع العامل الثابت)Fixed Factor

▪ اختر المتغير apt وضعه في مربع متغير التغاير Covariate

§ اختر الخيار Options لعرض الاحصاءات والاختبارات المصاحبة .

اختر المتغير method انقله في مربع العرض Display mean box وعلم باقي المربعات كما بالشكل التالي:

نافذة الخيارات في اختبار أنكوفا ببرنامج SPSS

3- انقر الزر Model وانقر الخيار Custom

انقل متغير التغاير apt و الطريقة method وعلمها معا ونقلهما للمربع الأيسر كما بالشكل التالي:

المخرجات:

Descriptive Statistics

Dependent Variable ACH

METHOD	Mean	Std Deviation	N
1.00	4.3333	1.5275	3
2.00	8.3333	1.5275	3
3.00	11.3333	1.1547	3
Total	8.0000	3.2787	9

حساب التفاعل

METHOD*APT.

Tests of Between-Subjects Effects

Dependent Variable: ACH

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	84.905[a]	5	16.981	46.513	.005	.987
Intercept	1.629	1	1.629	4.461	.125	.598
METHOD	1.416	2	.708	1.938	.289	.564
APT	10.343	1	10.343	28.337	.013	.904
METHOD * APT	.129	2	6.456E-02	.177	.846	.105
Error	1.095	3	.365			
Total	683.000	9				
Corrected Total	86.000	8				

a. R Squared = .987 (Adjusted R Squared = .966)

$$Partial\ \eta^2_{interaction} = \frac{Sum of\ Square_{interaction}}{Sum of\ Square_{interaction} + Sum of\ Square_{error}}$$

الاستخلاص:

The interaction is not significant, $F(2,3) = .177$, $p = .846$, partial eta square = .105.

التفاعل ليس دال إحصائيا مما يشير لتحقق شرط تجانس الإنحدار وبذلك يتحقق أحد شروط تحليل التغاير

وبعض البرامج الأخري غير برنامج SPSS تخرج جدول مستقل لتجانس الانحدار كما يلي :

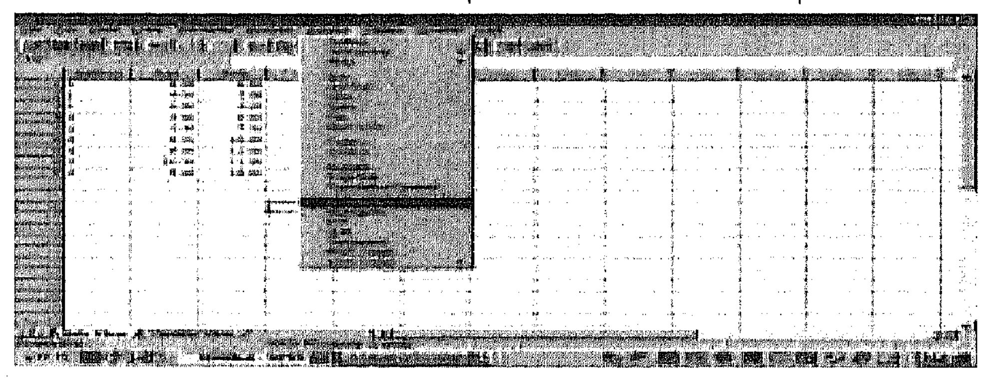
اختبار تجانس الانحدار

Source	SS	df	MS	F	P
between regressions	0.13	2	0.06	0.18	0.843671
remainder	1.1	3	0.37		
adjusted error	1.22	5			

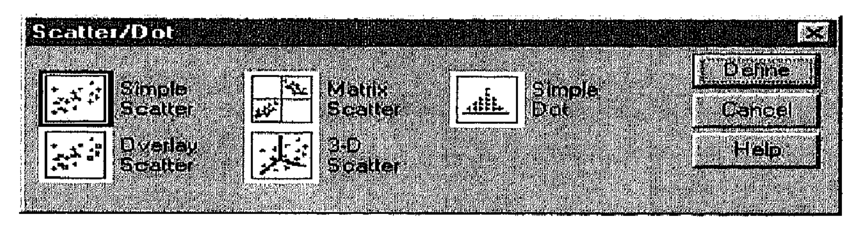
وهذا الجدول لا يظهر في برنامج SPSS وعدم دللالة قيمة F تعني تجانس الإنحدار

ختبار شرط العلاقة الخطية بين المتغير التابع – التحصل – ومتغير التغاير الاستعداد

من قائمة رسم Graph اختر Scatter ثم اختر النمط النشط البسيط



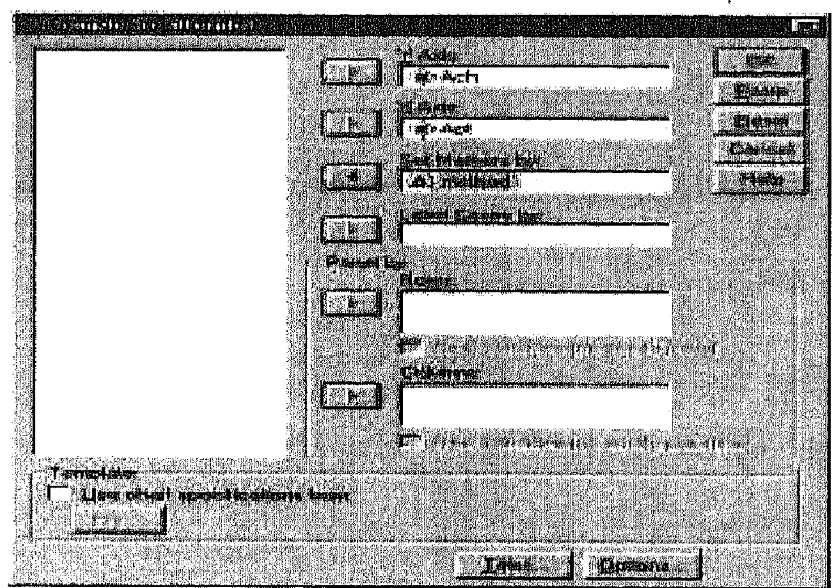
اختيار التشتت البسيط



انقر الزر Define وانقل المتغير Ach لمربع محور Y ثم ضع المتغير Apt في مربع محور X

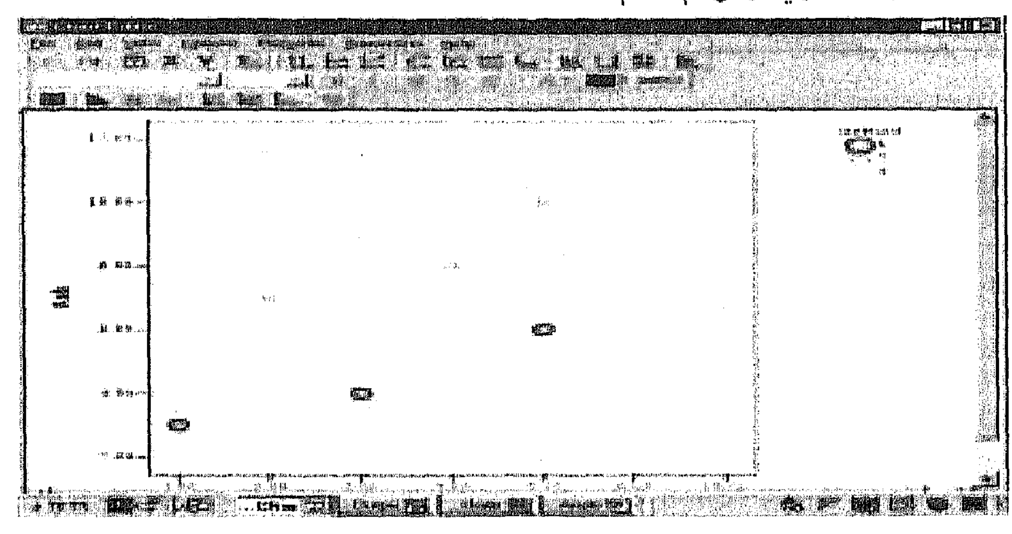
وانقل متغير العامل method لمربع Set Markers by كما بالشكل التالي :

نافذة اعدادات رسم الانحدار البسيط

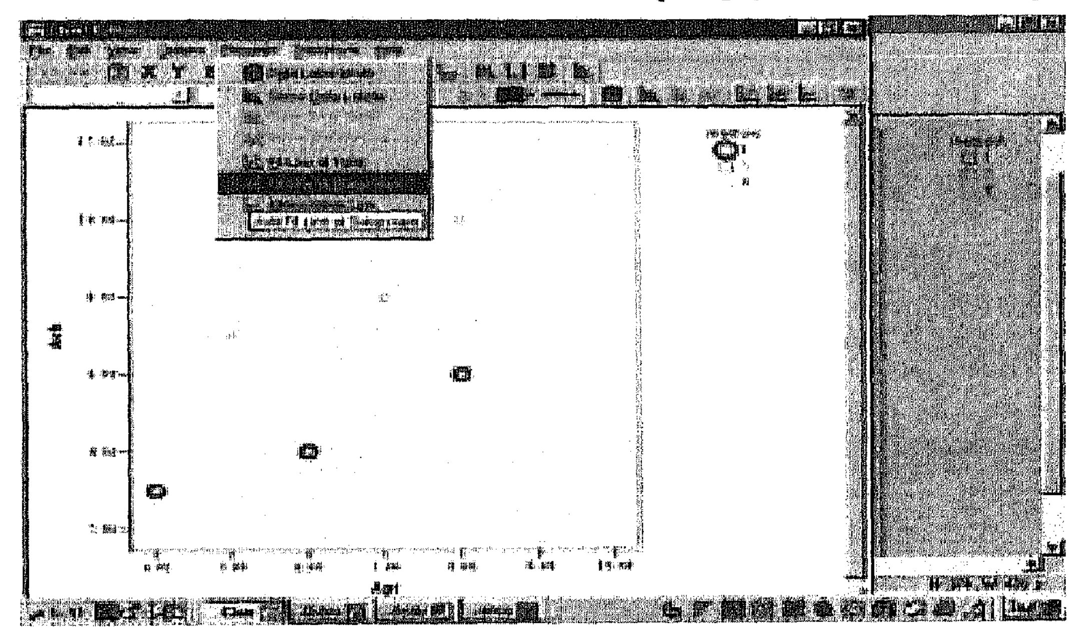


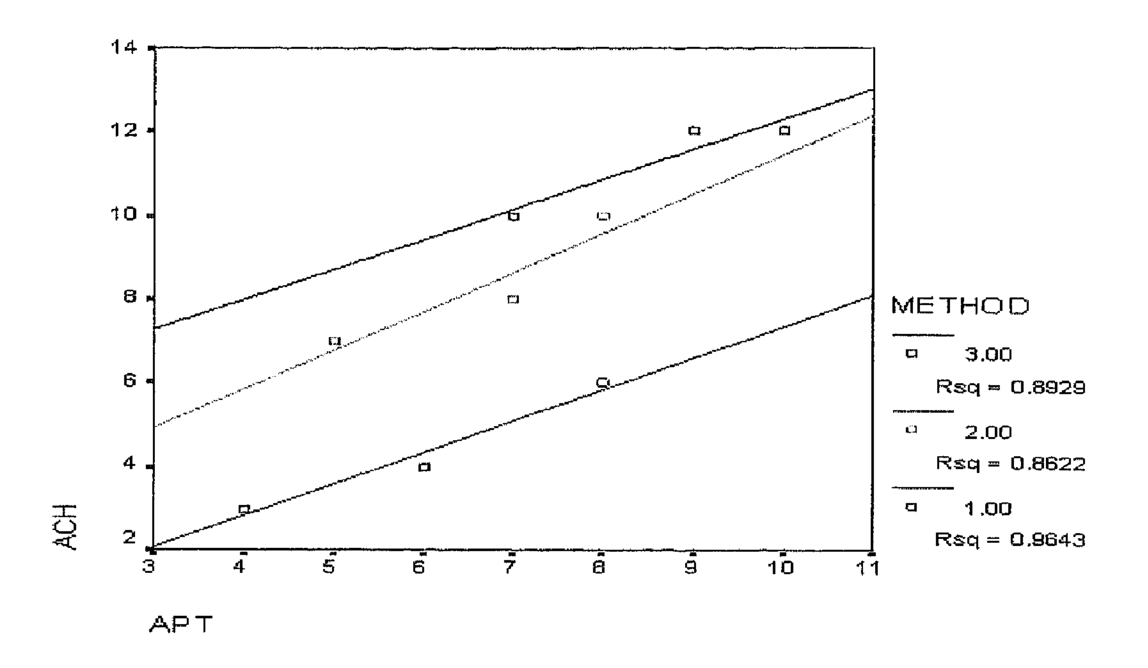
عدل المخطط للحصول علي خطوط الانحدار بنقر الرسم نقراً مزدوجاً

انقر دائرة الطريقة الأولي ومن ثم تعلم نقاطها



من قائمة elements اختر Fit subgroup





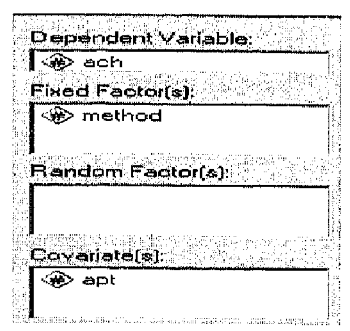
ومن الرسم يتضح أن متغير التغاير-الاستعداد الرياضي – يرتبط بعلاقة خطية مع المتغير التابع – التحصيل في الرياضيات – في الطرق الثلاث المستخدمة في التدريس ومن ثم يتحقق شرط تساوي الميل The assumption of equal slopes ومن ثم ندري اختبار تحليل التغاير انكوفا باتجاه واحد .

وهذا واضح من شبه التوازي في الطرق الثلاث و مربع معامل الارتباط مرتفع والمتقارب في القيمة

تطبيق اختبار أنكوفا

الفروق في التحصيل لا ترجع فقط للفروق عبر طرق التدريس الثلاث ولكن أيضا للفروق الابتدائية في الاستعداد ولكي نعزل أثر الاستعداد كما نراها في التطبيق القبلي لاختبار الاستعداد الرياضي .ومن ثم نستخدم اختبار أنكوفا الذي سوف نناقشه فيما يلي من خلال المثال ولكن أو التنويه إلي أن الخطوات السابقة ضرورية ويجب التوقف عن استخدام اختبار أنكوفا لدي عدم تحقق تلك الشروط وخاصة شرطي تجانس الانحدار وشرط الخطية الذي يشير ببساطة لارتباط تنبؤي بين متغير التغاير والمتغير التابع .وسوف نفترض انك أجريت تلك الاختبارات التحققية:

اختر Analyze \ General Linear Model \ Univariate وضع المتغيرات كما سبق ويوضحها الشكل التالي



اختر Model وعلم الزر Full factorial كما باشكل التالي :

Univariate: Model
Specify Model
Full factorial
Custom
Factors & Covariates:
Model:

انقر الزر Ok

المتوسطات المعدلة

Estimated Marginal Means

METHOD

Dependent Variable: ACH

METHOD	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	5.209[a]	.315	4.400	6.018
2.00	8.684[a]	.291	7.937	9.431
3.00	10.107[a]	.340	9.232	10.982

a. Evaluated at covariates appeared in the model: APT = 7.1111.

الآن يجـــب أن نختبر الفروق بين المتوسطات لنحسب الفرق بين الطرق الثلاث بعد عزل أثر الاستعداد الرياضي

Tests of Between-Subjects Effects

Dependent Variable: ACH

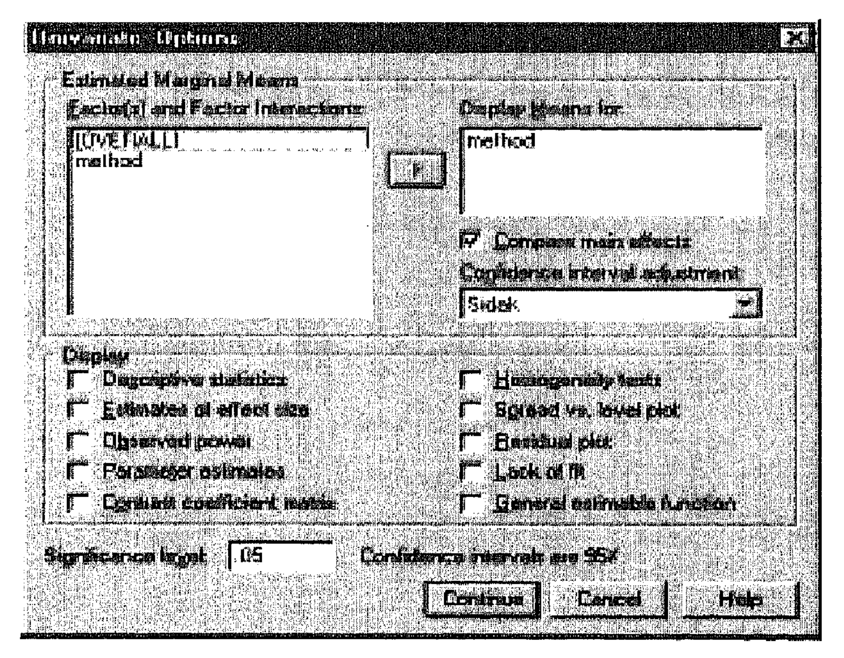
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	84.776[a]	3	28.259	115.401	.000	.986
Intercept	1.891	1	1.891	7.723	.039	.607
APT	10.776	1	10.776	44.005	.001	.898
METHOD	26.587	2	13.294	54.288	.000	.956
Error	1.224	5	.245			
Total	662.000	9				
Corrected Total	86.000	8				

a. R Squared = .986 (Adjusted R Squared = .977)

من الجدول ننظر لقيمة F المقابلة للطرق المستخدمة Method نجد أن قيمة F 54.288 وهـــي دالـــة عند مستوي أقل من 0.001 وبالتالي يجب أن نستخدم اختبار للمقارنة البعدية لـــتحديد أى الطـــرق تفـــوقت علي الطريقتين الأخريتين ، كما نلاحظ أن حجم التأثير كبير 0.956

$$Partial\ \eta^2_{method} = \frac{Sum of\ Square_{method}}{Sum of\ Square_{method} + Sum\ of\ Square_{error}}$$

ملحـــوظة: لـــن نتطرق لاستخدام الزر contrast للسهولة ولذلك من نافذة الخيارات السابقة



نتأكد من تعليم مربع الاختيار compare main effect وسوف نحصل علي نتائج من ضمنها الجدول التالي:

المقارنات البعدية باستخدام اختبار بونفروني

Pairwise Comparisons

Dependent Variable: Ach

(I) method	(J) method	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-3.474*	.412	.001	-4.929	-2.019
	3	-4.897*	.514	.001	-6.712	-3.083
2	1	3.474*	.412	.001	2.019	4.929
	3	-1.423	.469	.087	-3.080	.234
3	1	4.897*	.514	.001	3.083	6.712
	2	1.423	.469	.087	-.234	3.080

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

1- لا توجد فروق دالة إحصائيا عند مستوي 0.01 بين الطريقة 2 والطريقة 3 لصالح الطريقة 2 ونلاحظ ظهور علامة * مع المتوسطين

2- تـــوجد فروق دالة إحصائيا عند مستوي 0.01 بين الطريقة 1 والطريقة 3 لصالح الطريقة 1

3- تـــوجد فروق دالة إحصائيا عند مستوي 0.01 بين الطريقة 1 والطريقة 2 لصالح الطريقة 1

ملحوظة: اختبار أنكوفا بديل لاختبار أنوفا في حالة عدم تكافؤ المجموعات.

5-9 استخدام برمجية SPSS في حل المسائل.

مثال 9-1: يوجد 4 شعب في كل شعبة 10 طلاب درست بأربع طرق مختلفة ، والمطلوب معرفة إن كان هناك فرق معنوي في متوسط الشعب الأربعة عند مستوى معنوية 0.10

الفرضية:

هل تختلف متوسطات الشعب باختلاف طريقة التدريس عند مستوى معنوية 0.10 ؟

المتغيرات:

المتغير التابع Dependent : العلامة Mark

المتغير المستقل Independent : الشعبة Sectionn ولها اربع مستويات.

اذاً نستخدم تحليل التباين الاحادي One Way ANOVA

إدخال البيانات:

	sec	mark		sec	mark
1	1.00	5.00	21	3.00	7.00
2	1.00	6.00	22	3.00	5.00
3	1.00	3.00	23	3.00	6.00
4	1.00	2.00	24	3.00	8.00
5	1.00	4.00	25	3.00	5.00
6	1.00	10.00	26	3.00	10.00
7	1.00	7.00	27	3.00	7.00
8	1.00	3.00	28	3.00	3.00
9	1.00	4.00	29	3.00	4.00
10	1.00	7.00	30	3.00	6.00
11	2.00	8.00	31	4.00	10.00
12	2.00	7.00	32	4.00	8.00
13	2.00	7.00	33	4.00	9.00
14	2.00	9.00	34	4.00	9.00
15	2.00	2.00	35	4.00	4.00
16	2.00	9.00	36	4.00	9.00
17	2.00	8.00	37	4.00	9.00
18	2.00	4.00	38	4.00	5.00
19	2.00	5.00	39	4.00	6.00
20	2.00	8.00	40	4.00	7.00

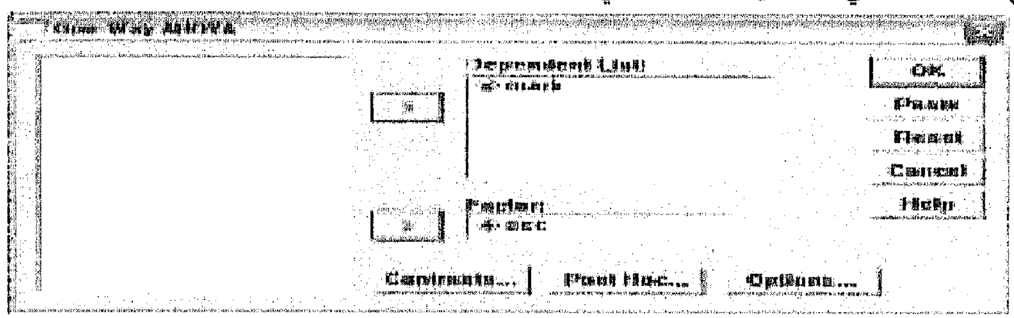
يستخدم اختبار ليفين للتأكد من تماثل التباينات Leven's homogeneity of variances test

لإجراء تحليل التباين نتبع الخطوات التالية:

Analyze – Compare Means – One-Way ANOVA...

ضع المتغير Mark في نافذة المتغيرات التابعة Dependent List:

ضع المتغير العاملي sec في نافذة عاملي Factor:



انقر مفتاح ...Options تظهر لك شاشة الحوار أدناه

اختر حساب الإحصاءات الوصفية Descriptive

اختر فحص التماثل بين المجموعات Homogeneity of variance test

ثم انقر زر Continue فتعود إلى شاشة حوار One-Way ANOVA

One-Way ANOVA: Options

Statistics
- [x] Descriptive
- [ ] Fixed and random effects
- [x] Homogeneity of variance test
- [ ] Brown-Forsythe
- [ ] Welch

- [ ] Means plot

Missing Values
- (•) Exclude cases analysis by analysis
- ( ) Exclude cases listwise

Continue | Cancel | Help

انقر زر الاختبارات البعدية | Post Hoc...

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

LSD | S-N-K | Waller-Duncan
Bonferroni | Tukey | Type I/Type II Error Ratio: 100
Sidak | Tukey's-b | Dunnett
Scheffe | Duncan | Control Category: Last
R-E-G-W F | Hochberg's GT2 | Test
R-E-G-W Q | Gabriel | 2-sided, < Control, > Control

Equal Variances Not Assumed

Tamhane's T2 | Dunnett's T3 | Games-Howell | Dunnett's C

Significance level: .05

Continue | Cancel | Help

من الاختبارات البعدية في حالة تساوي التباين اختر اختباري Scheffe و Tukey

من الاختبارات البعدية في حالة عدم تساوي التباين اختر اختباري Dunnett's C

ثم انقر زر Continue فتعود الى شاشة حوار One-Way ANOVA ثم انقر زر Ok

تظهر نتيجة تحليل التباين الأحادي كما هو مبين أدناه:

الإحصاءات الوصفية للمتغير التابع لكل فئة من فئات المتغير العاملي sec

Oneway

Descriptives

MARK

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	10	5.1000	2.42441	.76667	3.3657	6.8343	2.00	10.00
2.00	10	6.7000	2.31181	.73106	5.0462	8.3538	2.00	9.00
3.00	10	6.1000	2.02485	.64031	4.6515	7.5485	3.00	10.00
4.00	10	7.6000	2.01108	.63596	6.1614	9.0386	4.00	10.00
Total	40	6.3750	2.30593	.36460	5.6375	7.1125	2.00	10.00

نتائج تحليل التباين الاحادي: الاحصاءات الوصفية للمتغير التابع لكل فئة من فئات المتغير العاملي مثل المتوسطات الحسابيةMean، الانحرافات المعياريةStd. Deviation، والخطأ المعياري Std. Error، وفترات الثقة 95% Confidence Interval for Mean، وأقل قيمة Minimum واكبر قيمة Maximum للمتغير التابع لكل فئة من فئات المتغير العاملي.

Test of Homogeneity of Variances

MARK

Levene Statistic	df1	df2	Sig.
.205	3	36	.892

نـــتائج تحليل التباين الاحادي: اختبار ليفين لفحص تجانس التباين لفئات المتغير العاملي ينتج أن تباين المجموعات متساوية لأن قيمة مستوى الدلالة Sig=0.892 وهي اكبر من مستوى الدلالة ($\alpha = 0.05$).

ANOVA

MARK

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.075	3	11.025	2.277	.096
Within Groups	174.300	36	4.842		
Total	207.375	39			

نتائج تحليل التباين الأحادي: فحص فرضية الدراسة.

عدم وجود فروق ذات دلالة احصائية على مستوى أقل من ($\alpha = 0.05$)، حيث كانت قيمة مستوى الدلالة Sig أقل من 0.05

Post Hoc Tests اختبار الفروقات البعدية Multiple Comparisons

Dependent Variable: MARK

	(I) SEC	(J) SEC	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1.00	2.00	-1.6000	.98404	.377	-4.2502	1.0502
		3.00	-1.0000	.98404	.741	-3.6502	1.6502
		4.00	-2.5000	.98404	.070	-5.1502	.1502
	2.00	1.00	1.6000	.98404	.377	-1.0502	4.2502
		3.00	.6000	.98404	.928	-2.0502	3.2502
		4.00	-.9000	.98404	.797	-3.5502	1.7502
	3.00	1.00	1.0000	.98404	.741	-1.6502	3.6502
		2.00	-.6000	.98404	.928	-3.2502	2.0502
		4.00	-1.5000	.98404	.434	-4.1502	1.1502
	4.00	1.00	2.5000	.98404	.070	-.1502	5.1502
		2.00	.9000	.98404	.797	-1.7502	3.5502
		3.00	1.5000	.98404	.434	-1.1502	4.1502
Scheffe	1.00	2.00	-1.6000	.98404	.460	-4.4856	1.2856
		3.00	-1.0000	.98404	.793	-3.8856	1.8856
		4.00	-2.5000	.98404	.111	-5.3856	.3856
	2.00	1.00	1.6000	.98404	.460	-1.2856	4.4856
		3.00	.6000	.98404	.945	-2.2856	3.4856
		4.00	-.9000	.98404	.840	-3.7856	1.9856
	3.00	1.00	1.0000	.98404	.793	-1.8856	3.8856
		2.00	-.6000	.98404	.945	-3.4856	2.2856
		4.00	-1.5000	.98404	.516	-4.3856	1.3856
	4.00	1.00	2.5000	.98404	.111	-.3856	5.3856
		2.00	.9000	.98404	.840	-1.9856	3.7856
		3.00	1.5000	.98404	.516	-1.3856	4.3856
Dunnett C	1.00	2.00	-1.6000	1.05936		-4.9071	1.7071
		3.00	-1.0000	.96889		-4.1183	2.1183
		4.00	-2.5000	.99610		-5.6096	.6096
	2.00	1.00	1.6000	1.05936		-1.7071	4.9071
		3.00	.6000	.97183		-2.4338	3.6338
		4.00	-.9000	.96896		-3.9249	2.1249
	3.00	1.00	1.0000	.96889		-2.1183	4.1183
		2.00	-.6000	.97183		-3.6338	2.4338
		4.00	-1.5000	.90247		-4.3173	1.3173
	4.00	1.00	2.5000	.99610		-.6096	5.6096
		2.00	.9000	.96896		-2.1249	3.9249
		3.00	1.5000	.90247		-1.3173	4.3173

نـــتائج تحليل التباين الأحادي: اختبار Tukey واختبار Scheffe واختبار Dunnett C للفروقات البعدية.

ان الفرق بين متوسطات المجموعات 4،3،2،1 غير دال احصائياً، أي لا يوجد فروق ذات دلالة احصائية بين المجموعات الأربع.

Homogeneous Subsets

MARK

	SEC	N	Subset for alpha = .05
			1
Tukey HSD[a]	1.00	10	5.1000
	3.00	10	6.1000
	2.00	10	6.7000
	4.00	10	7.6000
	Sig.		.070
Scheffe[a]	1.00	10	5.1000
	3.00	10	6.1000
	2.00	10	6.7000
	4.00	10	7.6000
	Sig.		.111

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 10.000.

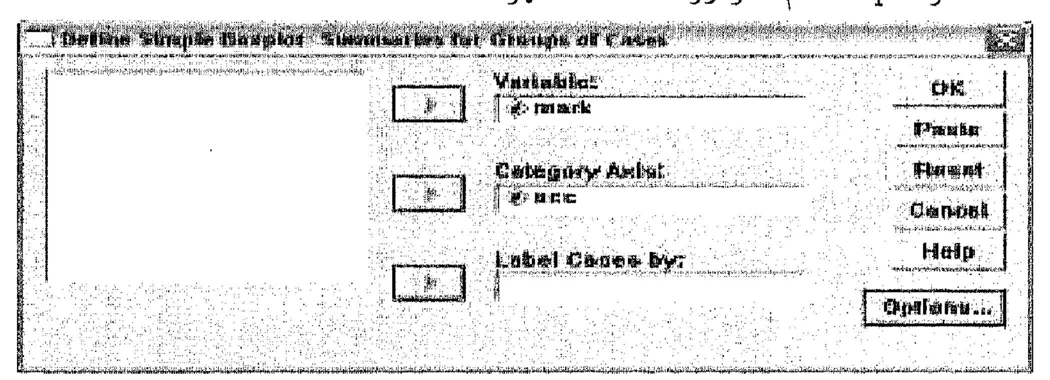
نـــتائج تحلـــيل التـــباين الأحادي: اختبار Scheffe واختبار Tukey للفروقات البعدية للمجموعات المتماثلة.

* استخدام الرسومات البيانية لتوضيح نتائج تحليل التباين الأحادي

ان توزيع التغير يختلف من فئة الى اخرى Graph – Boxplot...

Boxplot
Simple
Clustered
Define
Cancel
Help
Data In Chart Are
Summaries for groups of cases
Summaries of separate variables

انقر Simple ثم انقر زر Define تظهر الشاشة أدناه:



ضع المتغير mark في نافذة: Variable، وضع المتغير sec في نافذة: Category Axis:

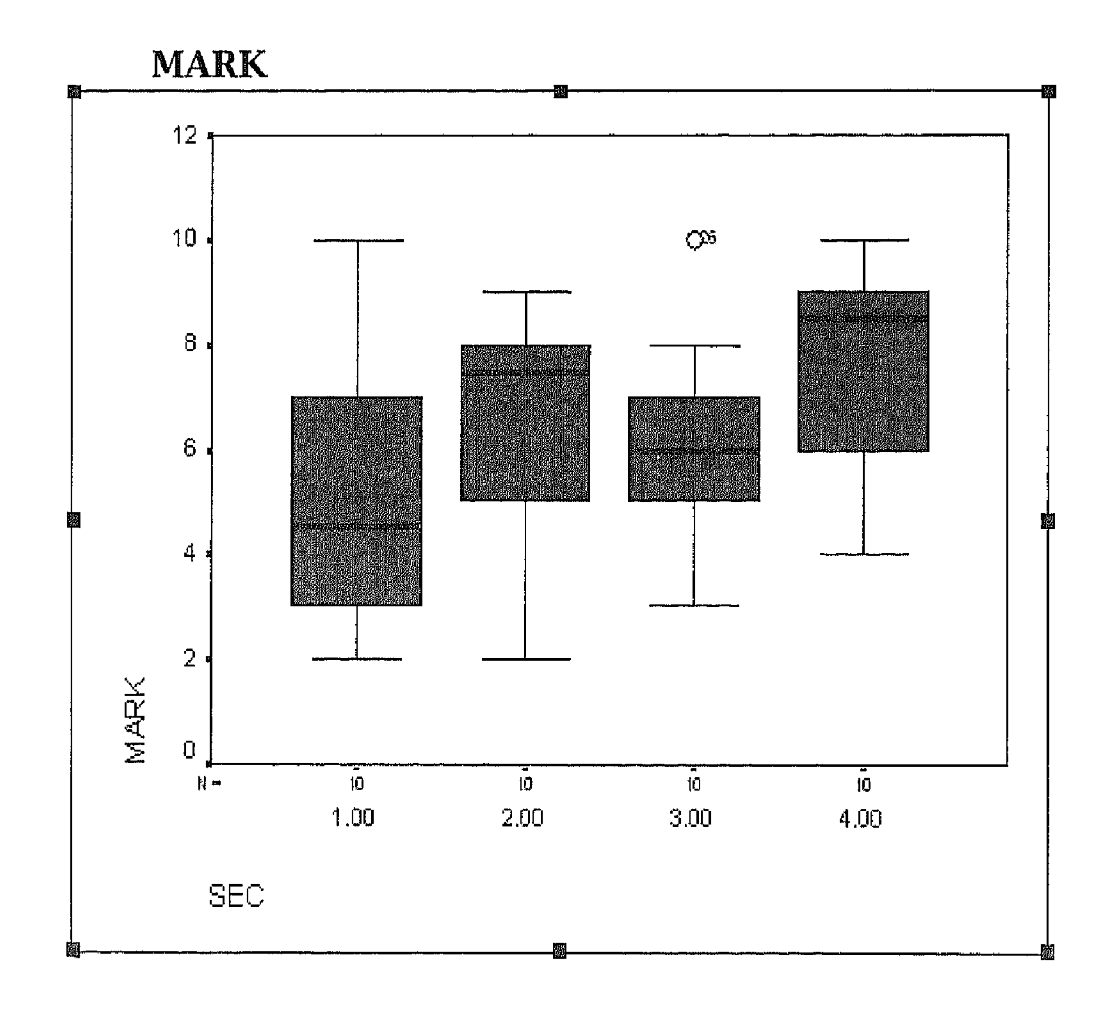
ثم انقر زر Ok ، تظهر لك شاشة المخرجات المبينة أدناه:

SEC

Case Processing Summary

	SEC	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
MARK	1.00	10	100.0%	0	.0%	10	100.0%
	2.00	10	100.0%	0	.0%	10	100.0%
	3.00	10	100.0%	0	.0%	10	100.0%
	4.00	10	100.0%	0	.0%	10	100.0%

MARK



6-9 تمـاريـــن Exercise.

س1: اجب عن ما يلي:

1- لماذا تكون قيمة F مساوية للواحد الصحيح ، عندما لا يكون هناك أثراً للمعالجة؟

2- وضـــح لمـــاذا نـــستخدم اسلوب تحليل التباين الأحادي بدلاً من تكرار استخدام اختبار t لاختبار الفروق بين ثلاثة أوساط أو أكثر.

3- علـــى افتراض ان رئيس جامعة مؤتة بصدد ترقية دكتور لدرجة أعلى، فما هو نوع الخطأ المتوقع الوقوع فيه إذا كان:

أ- الفرضية هي أن الدكتور يستحق الترقية، وتم قبول الفرضية الصفرية H0 خطئاً.

ب- الفرضية هي أن الدكتور يستحق الترقية، وتم رفض الفرضية الصفرية H0 خطئاً.

ج- الفرضـــية هي أن الدكتور يستحق الترقية، وتم قبول الفرضية الصفرية H0 بشكل صحيح.

د- الفرضـــية هي أن الدكتور يستحق الترقية، وتم رفض الفرضية الصفرية H0 بشكل صحيح.

س2- الجـــدول التالي يمثل نتائج ثلاثة شعب مختلفة في مادة القياس والتقويم في جامعة مؤته تم تدريسهم من قبل ثلاثة اساتذة مختلفين.

لنفرض الاساتذة كالتالي: الاستاذ 1 ، 2 ، 3

الشعبة: 1 ، 2 ، 3

رقم الطالب الشعبه	الاستاذ 1 الشعبه 1	الاستاذ 2 الشعبه 2	الاستاذ 3 الشعبه 3
1	45	66	60
2	50	78	50
3	55	90	70
4	60	67	90
5	35	75	95
6	50	72	93
7	62	66	88
8	64	70	74
9	71	80	51
10	80	72	35

1- هل هناك فروق معنوية بين متوسطات الشعب الثلاث، اوجد جدول تحليل التباين،

2- اختبر الفروق بمستوى معنوية ($\alpha = 0.05$).

3- اختبر تجانس التباينات.

4- هل يوجد هناك استاذ مميز ام لا.

س3- في دراسة لمعرفة أثر المعاملات الهرمونية علي إنتاج البيض أخذت عينة لأربعة معاملات هــــرمونية تتمثل في 10 دجاجات لكل معاملة وتم تسجيل إنتاج البيض السنوي لكل دجاجة كما هو موضح بعد:

المعاملــــة الأولى (X_1) : 180-190-200-210-230-210-220-180-240-240 بيضة.

المعاملــــة الثانية (X_2) : 230-260-250-270-280-290-230-240-280-270 بيضة.

المعاملــــة الثالثة (X_3) : 130-140-150-170-190-180-200-150-180-210 بيضة.

المعاملــــة الرابعة (X_4) : 240-250-230-210-260-220-240-260-220-270 بيضة.

والمطلوب :

1- وضع النموذج الإحصائي والنظرية الفرضية.

2- تكوين جدول تحليل التباين (ANOVA) .

3- مقارنة متوسطات المعاملات الهرمونية الأربعة.

س4- في تجربة لاستخدام أربعة أعلاف في تغذية أربعة مجموعات تتكون كل مجموعة من خمسة كتاكيت حيث كانت الزيادة في الوزن خلال فترة ما نتيجة هذه المعاملات كالتالي :

العلف	الزيادة اليومية في الوزن
1	52-21-42-49-55
2	63-98-30-112-61
3	92-95-81-97-42
4	154-85-169-137-169

والمطلوب:

1. وضع النموذج الإحصائي والنظرية الفرضية.
2. تكوين جدول تحليل التباين ANOVA
3. اختبار معنوية الفروق بين متوسطات المعاملات.

س5 : جـــربت 4 أنـــواع من هرمونات النمو علي أربع مجموعات من الفئران (كل مجموعة مولودة في بطن واحدة) لمدة معينة وكانت الزيادة في الوزن بعد هذه المدة هي:

هرمونات: النمو	الزيادة في الوزن بالجرام
أ	40-20-70
ب	60-40-60
جـ	50-40-80
د-	20-40-30

والمطلوب:

1- وضع النموذج الإحصائي والنظرية الفرضية.

2- تكوين جدول تحليل التباين (ANOVA) .

3- مقارنة متوسطات المعاملات الهرمونية الأربعة.

س6: اذا كان لديك العينات الثلاثة المبينة بالجدول ادناه:

الأولى العينة	الثانية العينة	الثالثة العينة
40 41 40.5 38.5	27 28 26.5 26.5	33 32 33.5 31.5
$X_1 = 40$ $S_1 = 1.08$	$X_2 = 27$ $S_2 = 0.71$	$X_3 = 32.5$ $S_3 = 0.91$

السؤال: هل في البيانات ما يكفي لوجود فرق بين المتوسطات؟ ولصالح أي المجموعات.

س7: في دراســـة لتأثير وجود الطلاب في الصفوف على تحصيلهم في مادة الإحصاء، قام أستاذ الإحـــصاء بأخذ عينات عشوائية ومستقلة من ثلاثة صفوف (يقوم تدريسها) كل منها مكـــون مـــن خمسة طلاب وقام الأستاذ برصد درجاتهم والجدول التالي يبينها. بمستوى معنوية $\alpha = 0.05$ اختبر ما إذا كان متوسط النتائج في اختبارات الأداء يختلف في تحصيل الطلاب.

الصف 1	**الصف 2**	**الصف 3**
58	96	66
62	87	65
77	66	88
90	55	92
80	78	60

1- هـــل هـــناك فـــروق معنوية بين متوسطات الصفوف الثلاث، اوجد جدول تحليل التباين. اختبر الفروق بمستوى معنوية ($\alpha = 0.05$).

2- اختبر تجانس التباينات.

3- هل يوجد هناك صف مميز ام لا.

س8: اراد باحث ان يعرف الفعالية النسبية لأربع طرق في الدعاية والاعلان فاختار 40 زبون ووزعهم عشوائياً في اربع مجموعات (10 في كل مجموعة). كان الاعلان بالطريقة الأولى عن طريق التلفزيون، والثانية عن طريق الراديو، والثالثة عن طريق الملصقات والرابعة عن

طـــريق الصحف ومن ثم اخذت الاراء عن المادة موضوع الاعلان فكانت نتائجها كما يلي:

الصحف	الملصقات	الراديو	التلفزيون
66	33	55	76
88	52	76	68
63	53	59	80
61	25	86	44
53	77	53	35
51	44	42	90
45	39	51	66
46	49	73	87
70	53	45	56

افحـــص فرضـــية الـــباحث حـــول وجـــود اختلافات بين الطرق الاربعة في الدعاية ($\alpha= 0.05$)

س9: قـــسمت مدينة اربد الى اربعة مناطق وهي الحي الشرقي والغربي والشمالي والجنوبي، وتم اختـــيار عيـــنة عشوائية تتكون من 10 بيوت معروضة للاجار في كل منطقة، وكانت الاجـــور المطلـــوبة لكل بيت بالدينار الاردني كما هو مبين في الجدول ادناه، والمطلوب معرفة ان كان هناك فرقاً معنوياً في معدل اجرة البيوت بين المناطق الاربعة، عند مستوى ($\alpha= 0.05$)

الحي الجنوبي	الحي الغربي	الحي الشرقي	الحي الشمالي
100	45	80	50
90	50	85	55
120	70	90	60
130	63	95	70
150	82	100	52
160	50	105	62
140	55	110	45
125	60	120	66
140	65	120	60
160	70	100	70

س10: اكمـــل الجدول التالي الذي يمثل خلاصة نتائج تحليل التباين لتجربة ما أجريت بهدف مقارنة أربعة مستويات معالجةن علماً بأن عدد الأفراد في كل مستوى (10).

Source of Variance مصدر التباين	Sum of Sequares (SS) مجموع المربعات	Degrees of freedom (DF) درجات الحرية	Mean. Sequares (MS) متوسط المربعات	F ف
Between Groups بين المجموعات	SSB		MSB=SSB/C-1 15	
Within Groups ضمن المجموعات	SSW 108			
Total الكلي	SST			

س11: في تجربة قامت بها أحدى المؤسسات الصحية لمعرفة أن كان هناك فرق في درجة الثقة بالـــنفس بين الأطفال المرضى والأطفال الأصحاء ، فأخذت عينة من الأطفال المرضى حجمهـــا $n_1 = 18$ وعيـــنة مـــن الأطفال الأصحاء حجمها $n_2 = 18$ فكانت نتائج التجربة تشير إلى حصول النقاط التالية:

	الوسط الحسابي	الانحراف المعياري
الأطفال المرضى	23.5	3.9
الأطفال الأصحاء	27.8	3.1

والمطلوب أجراء الاختبار عند مستوى معنوية $\alpha = 0.01$.

س12) لمعرفة الثقل المحوري للشاحنات المارة على طريقين تم إنشاؤهما حديثا لأحدى البلديات والمصممة لنفس المواصفات أخذت عينة ستكون من 31شاحنة من كل طريق واتضح بأن متوسط الحمولة لها والانحراف المعياري هي كالأتي:

	الوسط الحسابي	**الانحراف المعياري**
الطريق الأول	23.5	3.9
الطريق الثاني	27.8	3.1

فهـــل نـــستدل على وجود فروقات في حمولة الشاحنات المارة على كلا الطريقين عند مستوى معنوية $\alpha = 0.10$

س13: معمل فيه خطين إنتاجيين لإنتاج نوعين من المصابيح الكهربائية أخذت عينة من الخط الأول حجمها (60) مصباح فكانت نسبة المصابيح الغير صالحة 12% بينما في العينة الثانـــية الـــتي أخذت من الخط الثاني والتي كان حجمها (80) مصباح كانت نسبة المصابيح الغير صالحة 9%.

فإذا كان متوسط عمر المصابيح في عينة الأولى (900) ساعة والانحراف المعياري (20) وفي العينة الثانية كان المتوسط (970) ساعة والانحراف المعياري (17) . والمطلوب أيجاد:

أ- تقديـــر الفتـــرة للفـــرق بين متوسطي عمر المصابيح المنتجة في الخطين بمعامل ثقة (90%) .

ب- تقدير الفترة للفرق بين نسبتي المصابيح الصالحة في خطين الإنتاجين .

س14: اخـــتيرت عينة تتكون من 50 مدرسة إعدادية ، فكان معدل الدرجة النهائية للطالب لهذه المدارس هو 61 وبانحراف معياري مقداره 4.5 درجة . في حين أوضحت أمانة التربية المسؤوله عن هذه المدارس بأن المعدل النهائي يزيد على 62 وبانحراف معياري مقـــداره 5.1. والمطلـــوب اختبار مدى صحة ادعاء أمانة التربية عند مستوى معنوية $\alpha = 0.05$

س15: اختيرت عينتين من طلبة إحدى الجامعات ، الأولى تمثل الذكور والثاني إناث ، وطبقت عليهم إحدى أساليب التدريس الحديثة. وفي نهاية السن الدراسية كان عدد الناجحين لكل من العينتين هو كما يلي:

النوع	حجم العينة	عدد الناجحين
ذكور	86	50
إناث	60	45

فهـــل نستدل بأن هناك فرق بين الذكور والإناث في النتيجة النهائية عند مستوى منوية $\alpha = 0.01$

س16: في دراســة قامت بها إدارة التلفزيون لمعرفة إن كان برنامجها الترفيهي له نفس الأهمية لكافة الفئات العمرية ، فأخبرت عينة من المشاهدين ، وحصلت على النتائج المبينة في الجدول التالي . المطلوب اختبار أن كان هناك فروق بين الفئات العمرية اتجاه البرنامج المعنية عند مستوى معنوية $\alpha = 0.10$.

الرغبة	أقل من 18	18-39	40-59	60سنة فأكثر	المجموع
لا يرغب	32	18	17	11	78
يرغب	115	85	71	36	307
يرغب جدا	62	41	22	28	153
المجموع	209	144	110	75	538

س17: للمقارنة بين عدد الكيلو مترات التي يمكن قطعها بالغالون الواحد لنوعين من البنزين، تم اختــيار ســيارتي صالون من نفس النوع والموديل ، وتم تسيير كل منها عدة مرات فكانت النتيجة كالأتي:

عدد الكيلو مترات المقطوعة	
للنوع الأول من البنزين	للنوع الثاني من البنزين
78	70
80	81
76	79
81	72
85	69
78	75
82	66
83	70
86	68
90	75

والمطلوب اختبار ان كان هناك فرق بين تباين المسافات المقطوعة لكل من نوعي البنزين المستخدم عند مستوى معنوية $\alpha = 0.05$.

س**18**: المعطيات الآتية تمثل طول شعرة القطن (ملم) لخمسة أصناف من القطن لعينة عشوائية حجمهـــا 10 شعيرات من كل صنف . المطلوب اختبار الفرض القائل بأنه لا توجد فروق جوهرية بين هذه الأصناف الخمسة تحت مستوى معنوية $\alpha = 0.01$.

الصنف	طول شعرة القطن (ملم)										مجموع
A	39	39	38	37	36	37	39	39	38	39	381
B	39	40	38	39	38	41	41	38	40	38	392
C	35	36	36	37	35	36	35	35	38	37	360
D	36	40	38	39	39	41	42	41	40	40	396
E	35	36	35	34	33	32	33	32	32	34	336

س**19**: اختبر فرضية $H_0: \mu_1 = \mu_2 = \mu_3$ للمعطيات في الجدول الأتي :

الصنف	المشاهدات						المجموع
1	4	7	6	3			20
2	7	8	6	6	5	4	36
3	5	6	7				18

س**20**: استخدمت أربع طرق لأربع مجاميع من الطلبة لتعليمهم جدول الضرب، وكانت النتائج كمـــا يلـــي ، المطلوب اختبار فيما إذا كانت فروق جوهرية بين الطرق الأربع وعند المستوى معنوية $\alpha = 0.05$.

طريقة التعليم	المـــــشاهــــــدات						المجموع
1	7	6	8	5	9	7	42
2	8	9	10	7	8	6	48
3	7	8	10	5	6	3	39
4	8	6	5	4	9	4	36

س**21**: المعطيات في الجدول الأتي تمثل عاملين هما ويضم أربعة أنواع من الأمراض ويضم أربعة من المستشفيات والقيم تمثل عدد الأيام الذي يستغرقه المريض للعلاج في المرض.

المطلوب اختيار الفرضيات الآتية:

أ – هل هناك اختلاف في مدة العلاج بأختلاف المستشفى

ب– هل ان نوع المرض له علاقة بفترة العلاج ؟

ج– هل هناك تفاعل بين نوع المستشفى ونوع المرض ؟

العامل **B**	**1**	**2**	**3**	**4**	المجموع
العامل**A**					
	20	25	24	82	
	25	30	28	13	
1	22	29	24	62	534
	27	28	25	92	
	21	20	30	23	
	115	142	131	146	
	30	30	39	40	
	54	29	42	45	
2	30	31	36	50	765
	53	30	42	45	
	63	30	40	60	
	176	150	199	24	
	31	32	41	42	
	30	35	45	50	
3	40	30	40	40	766
	35	34	40	55	
	30	30	35	45	
	166	137	201	232	
	20	33	24	29	
	21	25	25	30	
4	20	28	30	28	509
	20	30	26	27	
	19	31	23	30	
	100	137	128	144	
المجموع الكلي	557	596	659	762	2574

س22: أجريت تجربة لبيان تاثير اربعة انواع في الاغذية في زيادة وزن مجموعة من الابقار تنتمي لـــثلاث سلالات مختلفة وتم اعطاء كل نوع في الابقار الاغذية الى خمسة ابقار في كل سلالة وكانت النتائج التالية التي تمثل مجموع الزيادة في وزن الابقار الخمسة لكل سلالة ولكل نوع من الغذاء والمطلوب:

1- تكوين جدول تحليل التباين
2- إجراء كافة الاختبارات الممكنة عند مستوى معنوية 0.05
3- في حالة وجود تاثير معنوية لنوع الغذاء على زيادة الوزن حدد أي الانواع كانت السبب.

السلالة	نوع الغذاء				
نوع	A	B	C	D	المجموع
1	91	98	112	109	408
2	113	116	114	119	162
3	127	121	116	121	185
المجموع	331	333	342	349	755

س23) في تجـــربة طبية معينة اخذت مجاميع من المرضى واعطيت جرعات بمستويات مختلفة من الـــدواء معين وبعد فترة زمنية معينة تم حساب نسب الشفاء للمجاميع وكانت كما يلي:

المستوى	نسب الشفاء						المجموع
A	81	80	75	80	70	70	456
B	70	70	68	65	69	70	412
C	62	63	64	67			256
D	68	66	65	63			262

1- اوجـــد جـــدول تحليل التباين للبيانات اعلاه واختبر فيما اذا كانت هناك فروق معنوية بين مستويات الجرعة عند مستوى معنوية (5%)
2- في حالة وجود فروق معنوية حدد ايا من مستويات الجرعة هي السبب في الفرق معنوي وعلق على النتيجة .

24: تمرين (17.9): اكمل الجدول تحليل تباين التالي واكتب النموذج الملائم ثم اختبر الفرضية القائله بوجود تاثير مشترك للصفوف والاعمدة بمستويّ معنوية $\alpha = 0.05$

مصدر تباين	درجة الحرية	مجموع المربعات	متوسطات المربعات	F
بين الاعمدة		12.25		1.306
بين الصفوف			167.25	53.52
تفاعل	9			
الخطأ				
الكلي	31	598		

الفصل العاشر

المقارنات المتعددة

Multiple Comparisons

10-1 مقدمة

10-2 أنواع المقارنات المتعددة

1- المقارنات المخطط لها

أ- طريقة المقارنات المتعامدة Orthogonal

ب- طريقة دن Dunn وتسمى ايضاً طريقة بنفوروني Bonferroni

2- المقارنات غير المخطط لها

أ- طريقة شافيه Scheffe

ب- طريقة توكي Tukey

ج- طريقة نيومان كولز Newman Kuelz

10-3 تمـاريـــن Exercise.

الفصل العاشر

المقارنات المتعددة
Multiple Comparisons

10-1 مقدمة

تحليل التباين Analysis of Variance: تعميم لاختبار t ويستخدم لاختبار ثلاث عينات أو أكثر، ويستخدم لمقارنة عدة متوسطات حسابية في آن واحد $H_0: \mu_1 = \mu_2 = ... = \mu_k$ ، وهو طريقة ذكية لاختبار اختلاف أوساط أكثر من مجموعتين دفعة واحدة من خلال التباين.

المقارنات البعدية Post hoc Comparisons: وهي عبارة عن اختبارات تتبعية نحتاج إليها عندما يوجد أكثر من وسطين وهذه المتوسطات مختلفة من الناحية الإحصائية (أي رفضت الفرضية الصفرية).

نلجأ للمقارنات البعدية عند رفض فرضية صفرية تتضمن ثلاثة متوسطات أو أكثر.

10-2 أنواع المقارنات المتعددة

يمكن تصنيف المقارنات المتعددة إلى نوعين رئيسيين هما:

1- المخطط له أو القبلي Planned or Apporiori ويتم الترتيب له بشكل مسبق قبل إجراء اختبار F وان استخدامه لا يتطلب ان تكون F ذات دلالة إحصائية.

2- غير المخطط له أو البعدي Post hoc or Apostroriori يتطلب ان تكون F ذات دلالة احصائية قبل إجراءه.

1- المقارنات المخطط لها

وتستخدم بدون اجراء تحليل التباين للبيانات وهي تعتمد على اجراء اختبار T-test ومن هذه الطرق ما يلي:

أ- طريقة المقارنات المتعامدة **Orthogonal**

من قائمة Analyze - Descriptive Statistics - Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
MARK * SEC	40	100.0%	0	.0%	40	100.0%

MARK * SEC Crosstabulation

			SEC				Total
			1.00	2.00	3.00	4.00	
MARK	2.00	Count	1	1			2
		% within MARK	50.0%	50.0%			100.0%
	3.00	Count	2		1		3
		% within MARK	66.7%		33.3%		100.0%
	4.00	Count	2	1	1	1	5
		% within MARK	40.0%	20.0%	20.0%	20.0%	100.0%
	5.00	Count	1	1	2	1	5
		% within MARK	20.0%	20.0%	40.0%	20.0%	100.0%
	6.00	Count	1		2	1	4
		% within MARK	25.0%		50.0%	25.0%	100.0%
	7.00	Count	2	2	2	1	7
		% within MARK	28.6%	28.6%	28.6%	14.3%	100.0%
	8.00	Count		3	1	1	5
		% within MARK		60.0%	20.0%	20.0%	100.0%
	9.00	Count		2		4	6
		% within MARK		33.3%		66.7%	100.0%
	10.00	Count	1		1	1	3
		% within MARK	33.3%		33.3%	33.3%	100.0%
Total		Count	10	10	10	10	40
		% within MARK	25.0%	25.0%	25.0%	25.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	21.429[a]	24	.613
Likelihood Ratio	26.693	24	.319
Linear-by-Linear Association	4.477	1	.034
N of Valid Cases	40		

a. 36 cells (100.0%) have expected count less than 5. The minimum expected count is .50.

بما ان قيمة مربع كاي = 21.429 وهذا يعني ان هناك علاقة قوية موجبة ولكن ليست ذات دلالة احصائية حيث ان مستوى المعنوية = 0.613

ب- طريقة دن Dunn وتسمى ايضاً طريقة بنفوروني Bonferroni

Oneway

ANOVA

MARK

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.075	3	11.025	2.277	.096
Within Groups	174.300	36	4.842		
Total	207.375	39			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: MARK
Bonferroni

(I) SEC	(J) SEC	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.6000	.98404	.676	-4.3474	1.1474
	3.00	-1.0000	.98404	1.000	-3.7474	1.7474
	4.00	-2.5000	.98404	.093	-5.2474	.2474
2.00	1.00	1.6000	.98404	.676	-1.1474	4.3474
	3.00	.6000	.98404	1.000	-2.1474	3.3474
	4.00	-.9000	.98404	1.000	-3.6474	1.8474
3.00	1.00	1.0000	.98404	1.000	-1.7474	3.7474
	2.00	-.6000	.98404	1.000	-3.3474	2.1474
	4.00	-1.5000	.98404	.817	-4.2474	1.2474
4.00	1.00	2.5000	.98404	.093	-.2474	5.2474
	2.00	.9000	.98404	1.000	-1.8474	3.6474
	3.00	1.5000	.98404	.817	-1.2474	4.2474

2- المقارنات غير المخطط لها

نعمل في البداية تحليل تباين، ثم نعمل مقارنات بعدية حسب القرار.

* الحصول على الفروق الحرجة:

هناك أكثر من آلية للحصول على الفروق الحرجة ولكننا سنهتم بثلاثة طرق وهي:

أ- طريقة شافيه Scheffe

ب- طريقة توكي Tukey

ج- طريقة نيومان كولز Newman Kuelz

أ- طريقة شافيه Scheffe

تستخدم هذه الطريقة في إجراء جميع المقارنات بين الاوساط وهي مفضلة على الطرق الأخرى عندما تكون حجوم الخلايا غير متساوية وفي اجراء المقارنات المعقدة، وحتى تكون المقارنة ذات دلالة احصائية فإنها يجب أن تكون أكبر من المقدار CDS

الفرق الحرج عند شافيه هو CDS (Critical Def. Scheffe)

$$CDS = \sqrt{(k-1)}\ \alpha\ F_{k-1,n-k}\ MSE\ (1/n_i + 1/n_j)$$

F: القيمة الحرجة عند مستوى الدلالة المعني

K: عدد المجموعات

Error with in main sequare :MSE

الاختلاف بين موقع وآخر فيما يتعلق بالفروق الحرجة يكمن فقط في اختلاف حجوم العينات، وفي حالة تساوي الحجوم للمجموعات المختلفة يكفي فرق حرج واحد.

$$CDS = \sqrt{2}(3.29)\ (1.83)(1/_{12} + 1/_{12}) = 1.42$$

أي فرق يزيد عن هذه القيمة الحرجة يكون دال احصائياً.

هناك فرق بين X'_1 و X'_3 دال احصائياً وهو لصالح X'_1 لأنها أكبر.

هناك فرق بين X'_2 و X'_3 دال احصائياً وهو لصالح X'_2 لأنها أكبر.

الفرق بين X'_1 و X'_2 غير دال احصائياً.

طريقة شافيه غير حساسة للفروق (متشددة)، بمعنى لا تعتمد الفرق أن يكون فرق إلا إذا كان ذا دلالة قوية، ولذلك نتوقع أن يكون الفرق الحرج عند شافيه أكثر منه في الطرق الأخرى.

ب- طريقة توكي Tukey أو المسماة Honestly Significant Difference (HSD)

يعطى الفرق الحرج في طريقة توكي Tukey حسب القانون التالي:

$$CDT = \alpha\, q_{k,v} \sqrt{MSE\,(1/2n_i + 1/2n_j)}$$

k: عدد المجموعات.

v: درجات الحرية للمقام.

$$CDT = 0.05\, q_{3,33} \sqrt{1.83(1/_{24} + 1/_{24})} = 3.49 \sqrt{1.83(2/_{24})} = 1.36$$

أي فرق أكثر من هذه القيمة الحرجة يكون دال احصائياً .

جميع الفروق دالة احصائياً.

ج- طريقة نيومان كولز Newman- Kuelz

وتفيد هذه الطريقة في المقارنات بين أزواج الأوساط فقط.

يعطى الفرق الحرج في طريقة نيومان كولز Newman Kuelz حسب القانون التالي:

$$CDNs = \alpha\, q_{c,v} \sqrt{MSE\,(1/2n_i + 1/2n_j)}$$

c: عدد المتوسطات ضمن مدى موضع الاهتمام (المقارنة)

v: درجات الحرية للمقام.

$X'_1\ X'_2\ X'_3$ تكون $c=3$ $X'_1 - X'_3$ ، $X'_1 - X'_2$ ، $X'_2 - X'_3$

ملاحظة: في حالة نيومان كولز نحتاج الى أكثر من فرق حرج حتى في حالة تساوي الحجوم.

بين $X'_2 - X'_3$ تكون $c=3$

نفس قيمة توكي $CDNs = CDT = 1.36$ لأن قيمة $k = c$

بين $X'_1 - X'_3$ تكون $c = 2$ وبين $X'_1 - X'_2$ تكون $c = 2$

$$CDNs = 0.05\, q_{2,33} \sqrt{1.83\,(1/_{24} + 1/_{24})} = 1.13$$

جميع الفروق دالة احصائياً.

مثال: في دراسة تشتمل على 5 مجموعات كانت النتائج كما في الجدول التالي:

المجموعة	X'	S^2	N
G1	50	48	20
G2	52	50	25
G3	55	51	25
G4	49	51	25
G5	44	50	25

1- هـــل تـــوجد فروق دالة احصائياً بين متوسطات المجموعات الخمس عند مستوى الدلالة الاحصائية $\alpha = 0.05$ ؟

2- هل يلزم مقارنات بعدية ؟

3- إذا لزم استخدم طريقة توكي وطريقة شافيه وطريقة نيومان كولز لتحديد فيما إذا كـــان الفرق بين متوسطات المجموعة الأولى والمجموعة الخامسة دال احصائياً عند مستوى الدلالة الاحصائية $\alpha = 0.05$ ؟

4- هـــل تـــوجد فروق دالة إحصائيا بين متوسطات المجموعات الخمس عند مستوى الدلالة الإحصائية $\alpha = 0.05$ ؟

المجموعة	X'	S^2	N	X' * n	SSB	SSW
G1	50	48	20	1000	0	912
G2	52	50	25	1300	100	1200
G3	55	51	25	1375	625	1224
G4	49	51	25	1225	25	1224
G5	44	50	25	1100	900	1200
المجموع			120	6000	1650	5760
X'_w	50	**SSB**	1650		**SSW**	5760

1- نجد الوسط الموزون

$$X' = \frac{(20*50)+(25*52)+(25*55)+(25*49)+(25*44)}{120} = 50$$

$$SSB = 20(50-50)^2+25(52-50)^2+25(55-50)^2+25(49-50)^2+25(44-50)^2 = 1650$$

$$SSW = 19(48)+24(50)+24(51)+24(51)+24(50) = 5760$$

Source of Variance مصدر التباين	Sum of Sequares (SS) مجموع المربعات	Degrees of freedom (DF) درجات الحرية	Mean. Sequares (MS) متوسط المربعات	F ف
Between Groups بين المجموعات	1650	5 - 1 4	MSB=SSB/C-1 =1650/4 = 412.5	MSB/MSW 412.5/50.09 8.24
Within Groups ضمن المجموعات	SSW 5760	N - C =120-5 = 119 115	MSW=SSW/N-C =5760/120-5 = 50.09	
Total الكلي	SST 7410	N - 1 =120-1 =119 119		

نجد قيمة F الحرجة $F_{0.05,4,115} = 2.45$

بما أن قيمة F المحسوبة 8.24 أكبر من > قيمة F الحرجة $F_{0.05,4,115} = 2.45$ لذلك نرفض الفرضية الصفرية. أي أن هناك فروق بين متوسطات المجموعات.

2- هل يلزم مقارنات بعدية ؟

بما أن هناك فروق بين متوسطات المجموعات إذا يلزم مقارنات بعدية.

3- إذا لـــزم استخدم طريقة توكي وطريقة شافيه وطريقة بيومان كولز لتحديد فيما إذا كان الفرق بين متوسطات المجموعة الأولى والمجموعة الخامسة دال احصائياً عند مستوى الدلالة الاحصائية $\alpha = 0.05$ ؟

جدول الفروق بين المتوسطات

	$X'_{5} = 44$	$X'_{4} = 49$	$X'_{1} = 50$	$X'_{2} = 52$	$X'_{3} = 55$
$X'_{5} = 44$	-	5	6	8	11
$X'_{4} = 49$		-	1	3	6
$X'_{1} = 50$			-	2	5
$X'_{2} = 52$				-	3
$X'_{3} = 55$					-

1- طريقة شافيه Scheffe

$CDS = \sqrt{(k-1)\ \alpha\ F_{k-1,n-k}\ MSE\ (1/n_i + 1/n_j)}$

$CDS = \sqrt{(5-1)\ (2.45)\ (50.09)\ (1/_{20} + 1/_{25})} = 6.25$

بما أن 6 أقل من < 6.65 إذا الفرق بين $X'1 - X'5$ غير دال احصائيا.

2- طريقة توكي Tukey

$CDT = \alpha\ q_{k,v} \sqrt{MSE\ (1/2n_i + 1/2n_j)}$

$CDT = 0.05\ q_{5,115} \sqrt{50.09\ (1/_{40} + 1/_{50})}$

$CDT = 3.92 \sqrt{50.09\ (1/_{40} + 1/_{50})} = 5.92$

بما أن 6 أكبر من < 5.92 إذا الفرق بين $X'1 - X'5$ دال احصائياً.

3- طريقة نيومان كولز Newman Kuelz

$CDNs = \alpha\ q_{c,v} \sqrt{MSE\ (1/2n_i + 1/2n_j)}$

$CDNs = 0.05\ q_{3,115} \sqrt{50.09(1/_{40}+ 1/_{50})} = 3.36\ (\ 1.51\ ) = 5.07$

بما أن 6 أكبر من < 5.07 إذا الفرق بين $X'1 - X'5$ دال احصائياً.

3-10 تمارين Exercise.

س1: متى نلجأ الى المقارنات البعدية؟

س2: اذكر انواع المقارنات البعدية.

س3: اشرح المقارنات المخطط لها

طريقة المقارنات المتعامدة Orthogonal

طريقة دن Dunn وتسمى ايضاً طريقة بنفوروني Bonferroni

س4: اشرح المقارنات غير المخطط لها

طريقة شافيه Scheffe

طريقة توكي Tukey

طريقة نيومان كولز Newman Kuelz

الفصل الحادي عشر

التحليل العاملي

Factor Analysis

الفصل الحادي عشر

التحليل العاملي

Factor Analysis

1-11 مقدمة Introduction

يعتمد التحليل الإحصائى بصفة عامة على نوع المشكلة محل القياس وخصائصها، ونوع البيانات سواء كانت إسمية أو ترتيبية، أو فترية أو نسبية، وأيضا على الأهداف المراد تحقيقها من الدراسة، وعلى ذلك فالتحليل الذى يستخدم لمشكلة معينة، قد لا يستخدم لمشكلة أخرى.

أصبح التحليل العاملى يحتل مكانة هامة فى البحوث بمختلف انواعها، حيث ان هذه العلوم تخضع لكثير من المتغيرات المتداخلة، التى يكون بينها مجموعة من الارتباطات السلبية أو الايجابية.

والتحليل العاملى أسلوب احصائى يساعد الباحث فى دراسة المتغيرات المختلفة بقصد ارجاعها الى اهم العوامل التى اثرت فيها ، فمن المعروف ان أى ظاهرة من الظواهر تنتج من عدة عوامل كثيرة ، وتعتبر الظاهرة محصلة لهذة العوامل جميعا (باهى وعبد الفتاح، 2006: 187).

أيضا فأن التحليل العاملى اسلوب احصائى يعمل على تجميع متغيرات ذات طبيعة واحدة فى تركيبة متجانسة مرتبطة داخليا فيما بينها فى تكوين يسمى عامل بحيث يرتبط كل متغير من هذه المتغيرات بهذا العامل ، اى ان كل متغير من هذه المتغيرات يتشبع على هذا العامل بقيم متفاوتة توضح الأهمية النسبية لكل متغير من هذه المتغيرات المرتبطة بالنسبة لهذا العامل (ابراهيم، 2002: 196).

ويشير عبد الخالق (1987) ان هناك اتجاها آخر عكس هذا الرأي وهو ان التحليل العاملى يقترح فروضا، وكلما نجح فى القيام بهذه المهمة انتهت وظيفة الوصف ليصبح جزءا من النظرية السيكولوجية من حيث هو احصاء يختصر العلاقات بين مجموعة من المتغيرات ويقترح علاقات سببية لم يسبق اكتشافها.

وان توليد الفروض ليس حكرا على التحليل العاملي فهو يشبه فى ذلك طريقة الملاحظة والعمـــل الاكلينيكـــى، الا ان الاخيرين يقلون عنه في درجة الدقة والصرامة وقد يسهل تكـــوين الفروض فى مجال تتوفر فيه ملاحظات كثيرة، الا ان اسهامات التحليل العاملى تصبح مهمـــة جـــدا فى المجـــالات الجديدة نسبيا، وذلك فى الاسراع فى تكوين فروض معقولة وقيمة واســتبعاد الفروض الضعيفة، ويتصل هذا الهدف باثبات الفروض او رفضها وبخاصة الفروض المتعلقة بتركيب الشخصية وتنظيمها كفروض الانماط والسمات ، مما يصعب احباطه او دحضه بالطرق غير العاملية. وبين مستوى اقتراح الفروض والتحقق منها متينة .

ومن خلال ما سبق نجد ان العوامل تكون :

- مفاهيم احصائية بحتة.
- مبادىء للتصنيف.
- وسيلة لاظهار العلاقات السببية (عبد الخالق ، 1987 : 100-101).

ويذكر ريتشارد وآخرون (1992) ان الغرض الجوهري من التحليل العاملى هو بقدر الامكان وصف علاقات التباين التلازمى بين العديد من المتغيرات بدلالة قليلة نسبيا مع متغيرات اخرى فى مجموعات مختلفة، ويمكن تصور ان هذه المجموعة من المتغيرات تمثل بنية اساسية متفردة تسمى عامل. (Johnson, Richard A. & Wichern , Dean .1992 : 296-297)

ويذكـــر جيلفورد (1961) انه اذا كان التحليل العاملى افضل اداة فعالة لاستخراج المعلـــومات من البيانات فانة ينبغى ان نكون على علم بان التحليل العاملى ليس لة قوة سحرية تكـــشف عن تلك المعلومات التى تتضمنها البيانات المتجمعة ودائما ينبغى على كل من يحاول اســـتخدامة الكـــشف عـــن معلـــومات سيكولوجية تتعلق بالتنظيم العقلى المعرفى أو بسمات الشخصية ان يبدأ بفروض واضحة قابلة للاختبار (Guilford, J.P., 1961 : 231)

ونظـــرا لـــتعدد أهداف البحث وتعدد متغيراته الاجتماعية والاقتصادية، فإنه تم اختيار الاساليب الإحصائية التى تتفق مع طبيعة البيانات وتسمح بتحقيق الهدف الأول من البحث وهو تحديـــد الهيكل البنائي الاجتماعي والاقتصادي هي أساليب التحليل متعدد المتغيرات وهى أكثر الأســـاليب ملاءمـــة للمتغيرات والعلاقات المحددة لتكوين هذا الهيكل، أما أساليب التحليل التى تحقـــق الهدف الثانى من البحث وهى إعادة توزيع عبء نظام التأمين الاجتماعي فهي النماذج

الرياضية المستخدمة في بحوث العمليات والتي تركز على إعادة توزيع الموارد وهي أكثر ملاءمة لتحقيق التوزيع الأمثل لهذا العبء.

التحليل العاملي هو أسلوب إحصائي متعدد المتغيرات، يسعى إلى تحديد الأبعاد أو العوامل التي تساعد في وصف ظاهرة معقدة، عن طريق تحليل مصفوفة الارتباط (معاملات الارتباط البسيطة) بين المتغيرات المختلفة الداخلة في وصف الظاهرة، وصولاً إلى عوامل Factors محددة تكمن وراء طبيعة العلاقات الداخلية بين مجموعة المتغيرات في هذه الدراسة. وفي سبيل ذلك يسعى التحليل العاملي إلى تقليل البيانات Data Reduction بتحديد عدد العوامل القليلة التي تفسر معظم التباين في عدد كبير من المتغيرات، فبدلاً من أن يكون لدينا (35) متغيراً مثلاً يختصر التحليل العاملي هذا العدد إلى (06) عوامل مثلاً. وعادة ما تكون البيانات هي قيم (درجات أفراد على متغيرات نفسية أو اجتماعية أو تربوية).

يمثل أسلوب التحليل العاملي خطاً (نمطاً) من البحث مختلف تماماً عن الطرق الرياضية الأخرى في العلوم الاجتماعية، لاعتماده على افتراضات إحصائية، وهي نظرية شائعة ومفضلة لدى كثيرين، لأنها تحاول أن تجيب على السؤال الذي طالما سأله العلم: ما هو أقل عدد من المفاهيم التي يمكن أن تنظم تعقد الظاهرة وتصفها؟ هيم (عبدالخالق، 1994).

يظهر من هذا أننا نستطيع أن نستخدم هذا الأسلوب الإحصائي في تنظيم مجال جديد يحتاج للتعرف على خصائصه ومتغيراته، وهي حاجة يسعى إليها الباحث عندما يطرق مجالاً جديداً لا يعرف كل متغيراته أو مدى تعلق المتغيرات المختلفة بظواهره الرئيسية، والنتيجة المباشرة لهذه الخطوة الاستكشافية هي إعادة الدراسة والتناول للمتغيرات المهمة في المجال، وبناء الفروض التي تفسر العلاقات بين هذه المتغيرات. وتستند فلسفة التحليل العاملي إلى تحليل الارتباطات بين المتغيرات، بغرض استخلاص أقل عدد ممكن من العوامل التي تعبر عن أكبر قدر من التباين بين المتغيرات. وبذلك يبدأ التحليل العاملي بحساب معاملات الارتباط البسيطة بين المتغيرات وتسجيلها في مصفوفة تصلح لهذا النوع من التحليل، ويكون الهدف هو توضيح وتفسير العلاقات بين تلك المتغيرات، وينتج عنها عدد قليل من المتغيرات الجديدة (المفترضة) تسمى بالعوامل التي تحتوي على كل المعلومات الأساسية. وينتهي التحليل بمصفوفة عوامل ما قبل التدوير ومصفوفة عوامل بعد التدوير، تلك العوامل التي أدت إلى ذلك الارتباط.

ويمكن القول بأن التحليل العاملي نشأ في كنف علم النفس، حيث البدايات الأولى على يــد الرواد الأوائل لعلم النفس من أمثال ثورنديك Thorndike وبيرسون Pearson وهوتلنج Hotteling وطومــسون Tomson وجيلفورد Gulford وجالتون Galton وهولزنجر K.J. Holzingerوبــيرت C. Burt وثيرستون L.L. Thurston وإلكسندر... W.P Alexander إلخ. ثم انتقل إلى التطبيقات العملية والعلمية في شتى فروع المعرفة (باهي، 2002).

ويــرجع الفــضل في ذلك إلى سبيرمان C. Spearman منذ عام 1863م، بأن طور أفكــاره وأضــاف أبعاداً جديدة للمفهوم في دراسته التي نشرها عام 1904م، حيث بين أن العامــل هــو السبب في الارتباط الموجب بين أي ظاهرتين، وفي تطور لاحق أعلن سبيرمان أن العامــل هــو الــسبب المباشر لوجود الارتباطات الموجبة القائمة بين أي عدد من المتغيرات أو المقاييس. وفرق سبيرمان بين عاملين هما: العامل العامGeneral Factor ، وهو العامل المشترك بــين جميع المتغيرات، والعامل الخاص Specific Factor، وهو الذي يميز النواحي الخاصة التي ينفرد بها المتغير عن غيره من المتغيرات الأخرى .ولذا فمعامل ارتباط أي عاملين خاصين يساوي الــصفر. ولــذلك سميت نظرية سبيرمان العاملية بنظرية العاملين، وقد عدل بعض العلماء، مثل هولنجز، نظرية العاملين فأضاف لها نوعاص من العوامل التي توجد في طائفة من المتغيرات دون غيرهــا، وسماهــا بالعوامل الطائفية .Group Factor وطبقاً لهذه النظريات نجد أنه من الممكن تــصنيف العوامل التي يتوصل إليها الباحثون في البحوث النفسية والتربوية والاجتماعية إلى ثلاثة أنواع هي:

1- **العامــل العام**: هو العامل المشترك الذي يوجد في جميع المتغيرات) الاختبارات) التي تخــضع للتحليل العاملي، ويعبر عنه في هذه الحالة بالنمط العام، كما هو الحال في الذكاء العام على سبيل المثال.

2- **العامل الطائفي**: هو العامل الذي يوجد في بعض المتغيرات (الاختبارات) التي تخضع للتحليل، وليس فيها كلها، وهو يفسر ارتفاع قيم معاملات الارتباط بين الاختبارات التي تقيس الــذكاء، ومن أمثلة العوامل الطائفية القدرات العقلية الموجودة في الذكاء على سبيل المثال، مثل القــدرة المكانــية أو القــدرة الاســتدلالية، والتي تتسم بأنها عوامل ضيقة وغير قابلة لإعادة الاستخراج.

3- العامـــل الخاص أو النوعي: هو العامل الذي يختص بنوع واحد من أنواع السلوك الإنـــساني، ويـــوجد في متغير (اختبار) واحد فقط، أو عدة اختبارات تعكس جميعاً نفس المتغير المقـــاس، كاختبار الحساب، أو اختبار معاني الكلمات والتشابهات...إلخ .والتمييز بين العوامل الثلاثة (العام والطائفي والنوعي) ليس تمييزاً قاطعاً، إنما يتوقف على عدد المتغيرات (الاختبارات) الخاضـــعة للتحلـــيل، ومـــدى تباين أو تجانس استجابات أفراد العينة التي يجرى عليها التحليل العاملـــي، عـــدد أفـــراد العيـــنة التي يجرى عليها التحليل العاملي، مدى تجانس أو تباين هذه الاختبارات في قياسها لما تقيس، حجم أو قيم معاملات الارتباطات البينية للاخبارات الخاضعة للتحليل العاملي.

.بمعنى أن العامل الذي يظهر في مجموعة أقل عدداً من الاختبارات على أنه عامل عام، قد يظهـــر كعامل طائفي في مجموعة أكبر عدداً من الاختبارات، خاصة إذا مالت مجموعة منها إلى التجانس (الزيات، 1995).

2-11 مفهوم التحليل العاملي Factor Analysis Concept

التحلـــيل العاملـــي (Factor Analysis) هـــو أسلوب إحصائى يستهدف تفسير معـــاملات الارتباطات الموجبة التي لها دلالة احصائية – بين مختلف المتغيرات، وبمعنى آخر فإن التحلـــيل العاملـــي عملية رياضية تستهدف تبسيط الارتباطات بين مختلف المتغيرات الداخلة فى التحلـــيل وصـــولا إلى العوامل المشتركة التى تصف العلاقة بين هذه المتغيرات وتفسيرها. ويعد التحلـــيل العاملي منهجاً إحصائياً لتحليل بيانات متعددة ارتبطت فيما بينها بدرجات مختلفة من الارتـــباط التلخـــصي فى صورة تصنيفات مستقلة قائمة على أسس نوعية للتصنيف ، ويتولى الـــباحث فحص هذه الأسس التصنيفية واستشفاف ما بينها من خصائص مشتركة وفقاً للاطار النظري والمنطق العلمي الذى بدأ به، حيث يبدأ التحليل العاملي ، بحساب الارتباطات بين عدد مـــن المـــتغيرات مثل أ ، ب، جـــ ، د ، هـــ، و، مثل الذكاء ، القلق ، الانطواء، التحصيل، والاكتـــئاب مـــثلا، ونحصل على مصفوفة من الارتباطات بين هذه المتغيرات لدى عينة ما ، ثم نـــتقدم بعـــد ذلك لتحليل هذه المصفوفة الارتباطية تحليلا عاملياً لنصل إلى أقل عدد ممكن من المحاور أو العوامل تمكننا من التعبير عن أكبر قدر من التباين بين هذه المتغيرات ، ذلك أن توقفنا

عـــند فحص هذه المصفوفة الارتباطية التي تتكون من عشرة معاملات ارتباط لا يؤدى إلى فهم كامـــل للمجـــال المـــشترك فيما بينها جميعا، حيث يبين كل معامل من معاملات الارتباط فى المصفوفة علاقة بسيطة بين متغيرين فقط من متغيراتها دون أن ينبئ بأهمية أو دور هذه العلاقة بين هذين المتغيرين ومتغير ثالث ، وعلى ذلك لا نستطيع عند هذا المستوى أن نصل لتقدير للعلاقة المـــشتركة بين ثلاثة متغيرات معاً أو بين متغيرات المصفوفة الخمس إذ أن حصولنا على معامل للارتباط بين أ ، ب قدره 0.7 ومعامل آخر بين ب ، جـ قدره 0.7 أيضا لا يعنى بالضرورة أن الارتباط بين أ ، جـ يساوى 0.7 كذلك فقد يكون ما هو مشترك بين أ ، ب غير ما هو مـــشترك بـــين ب ، جـــ ، ولا تصلح العلاقة الثنائية بين ب وأي من المتغيرين أ ، جـــ لتقدير العلاقة بينهما فى معاملات الارتباط البسيطة (فرج، 1991).

يهـــدف أســـلوب التحلـــيل العاملي إلى تلخيص المتغيرات المتعددة في عدد أقل تسمى (عوامل) بحيث يكون لكل عامل من هذه العوامل دالة تربطه ببعض (أو كل) هذه المتغيرات. ويمكـــن من خلال هذه الدالة إعطاء تفسير لهذا العامل بحسب المتغيرات التي ترتبط معه بشكل قـــوي. ولقد نشأ هذا الأسلوب أساساً من أجل تحليل التجارب والمقاييس النفسية بحيث يمكن إرجاع مجموعة معينة من الاختبارات إلى عامل الذكاء وأخرى إلى عامل الذاكرة وهكذا، وإن كان هذا لا يعني أن هذا الأسلوب لا يستخدم في مجالات أخرى. وترتكز فكرة التحليل العاملي علـــى استخلاص مجموعة من العوامل مرتبطة بالمتغيرات الأصلية، بحيث تفسر هذه العوامل أكبر نـــسبة ممكنة من التباين في المتغيرات الأصلية. ويمكن استخدام التحليل العاملي لتحويل مجموعة مـــرتبطة من المتغيرات إلى مجموعة أخرى مستقلة تربطها بالمجموعة الأولى علاقات خطية. وفي كل الأحوال تمثل العلاقة بين المتغيرات الأصلية والعوامل في شكل معادلات على النحو التالي:

مثال على المعادلة $F_1 = \alpha_{11} X_1 + \alpha_{12} X_2 + \dots + \alpha_{1n} X_n$

أما المعادلات كاملة فهي كالتالي :

$$F_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \ldots + \alpha_{1n}X_n$$

$$F_2 = \alpha_{21}X_1 + \alpha_{21}X_2 + \ldots + \alpha_{2n}X_n$$

$$\vdots$$

$$F_m = \alpha_{m1}X_1 + \alpha_{m2}X_2 + \ldots + \alpha_{mn}X_n$$

3-11 أهمية التحليل العاملي وميادينه :

يمكن تطبيق أسلوب التحليل العاملي بنجاح في عدد كبير من الميادين العلمية، واختصار الوقت والجهد اللازمين للتحليل في العديد من الأبحاث، ويمكن إيجاز أهم هذه التطبيقات فيما يلي:

في مجال الإحصاء:

يُعتمد على التحليل العاملي في دراسة الارتباط والانحدار المتعدد بطريقة سريعة ودقيقة. فعلى سبيل المثال يستخدم التحليل العاملي لإثارة عدد من الفروض التي لها علاقة بالعوامل السببية، أو يستخدم لفحص المتغيرات قبل استخدامها في تحليل آخر، مثلاً: يمكن استخدام التحليل العاملي للبحث عن العلاقات الخطية المتعددة بين المتغيرات Multiconltinerity قبل تطبيق الانحدار المتعدد وتحويلها إلى عوامل مستقلة عن بعضها.

في مجال العلوم النفسية والاجتماعية والتربوية:

كُثر استخدام التحليل العاملي في هذا المجال، وذلك في تحليل النشاط العقلي المعرفي إلى قدراته المختلفة وتحليل النواحي المزاجية للشخصية إلى سماتها المتعددة وتحليل الاتجاهات والقيم الاجتماعية والميول المهنية.

في مجال بناء الاختبارات :

يعتمد بُناة الاختبارات الحديثة في دراسة مفردات الاختبارات على معرفة المكونات الرئيسية للظواهر التي تخضعها للقياس، ويُعد التحليل العاملي أدق وسيلة لمعرفة صدق هذه المكونات لقياس الظاهرة، وهو ما يسمى "بالصدق العاملي".

في مجال العلوم السياسية والتجارية :

حيث أن التحليل العاملي يقوم على الإيجاز الدقيق، فقد استخدم بنجاح كبير في دراسة الظواهر المعقدة التي تتأثر بعدد كبير من المؤثرات والعوامل المختلفة كالعلوم السياسية والإدارية، ودراسة العوامل المؤثرة في أسعار السلع والعملات وأجور العمال وما إلى ذلك.

مجالات أخرى كثيرة مثل مجال العلوم الطبية، ومجال العلوم الطبيعية :

يستخلص من ذلك أن التحليل العاملي ليس وقفاً على علم النفس أو التربية فقط، ولكنه أسلوب علمي إحصائي من أساليب الدراسة التحليلية التي تهدف إلى التقسيم والتبويب والتصنيف لجميع القوى والمؤثرات الفعالة في ظاهرة معينة (باهي 2002م، غنيم 2000، المالكي 2000).

4-11 أهداف التحليل العاملي :

كما أن الهدف الأساسي من التحليل العاملي هو وصف علاقات التغاير بين عدد كبير من المتغيرات بدلالة عدد قليل من المقادير غير المشاهدة التي تسمى العوامل، ويعتمد النموذج العاملي أساساً على الفكرة التالية: افتراض إمكانية تجميع المتغيرات بناءً على معاملات الارتباط بينها، وهذا يعني أن جميع المتغيرات الموجودة في مجموعة معينة مرتبطة مع بعضها ارتباطاً قوياً، ولكن ارتباطها بمتغيرات المجموعات الأخرى ارتباط ضعيف، ومن الممكن أن نتصور هنا أن كل مجموعة من المتغيرات تمثل عاملاً واحداً، وهو المسئول عن الارتباط المشاهد بينها (باهي 2002م، غنيم 2000، المالكي 2000).

ويسعى أسلوب التحليل العاملي إلى استخلاص العوامل من المتغيرات بحيث:

1. يكون العامل الأول هو أكثرها ارتباطًا بالمتغيرات أو أكثرها تفسيًرا للتباين المشترك يليه العامل الثاني وهكذا.
2. أن يكون في كل عامل عدد غير قليل من المعاملات الصفرية.
3. أن يسهل تفسير هذه العوامل على ضوء علاقاتها بالمتغيرات.

من أهم أهداف العلم تنظيم الحقائق والمفهومات تنظيما بوضع ما بينها من علاقات، أو تقسيمها على أساس ما بينها من أوجه التشابه والاختلاف والتحليل العاملي وسيلة من وسائل

التبسيط العلمي والتقسيم العلمي ويذكر "كاتل" (Cattell, 1952, 11) أن هدف المنهج العلمي اكتشاف الحقائق والعلاقة بين هذه الحقائق، ولأهداف عملية، واكتشاف القوانين التنبؤية، ويضيف أن التحليل العاملي منهج كلي يهدف إلى اكتشاف العموميات الاساسية، الوظيفية والعضوية ، بدلا من أن ينوه البحث فى عدد ضخم من المتغيرات التى تعد كالذرات، ولذلك يقترح " كاتل " أن يسمى بالتركيب العاملي أو على الأقل بتركيب المتغيرات. وبمعنى أضيق يحدد " سولمون دياموند" (عبد الخالق ، 1994 : 99). أهداف التحليل العاملي بأنه تكوين الفروض واختبارها ، وتحديد أصغر عدد من العوامل المحددة التى يمكن أن تفسر العلاقات التى نلاحظها بين عدد كبير من الظواهر الواقعية وإلى أي مدى يؤثر كل من هذه العوامل فى كل متغير ؟ أن أوضح وظيفة للتحليل العاملي تتمثل في خفض أو اختزال مكونات جداول الارتباطات إلى اقل عدد ممكن ليسهل تفسيرها .

لقد بين " أيزنك (Eyzanck, 1953) " أن للتحليل العاملي ثلاثة أهداف أساسية يروم تحقيقها، ويرتبط بهذه الأهداف ثلاث وجهات للنظر إلى طبيعة العوامل، وعدد كبير من طرق استخراج العوامل والتدوير ، وهى الأهداف ذاتها لأي فرع من فروع الاحصاء وهى:

1. الوصف.
2. البرهنة على الفروض.
3. اقتراح فروض من البيانات الأولية.

ومعظم علماء النفس يدركون هذه الاستخدامات الثلاثة للاحصاء، ولكن تظهر هذه المشكلة عندما تنطبق هذه الاهداف على التحليل العاملي، ويناقش " أيزنك " استخدامات التحليل العاملي على هذه المستويات الثلاثة ، مع تعريف العامل فى كل مستوى. فبالنسبة للهدف الأول فإن العامل احصاء مختصر يهدف إلى اقتصاد فى الوصف، ويصف علاقات مستقيمة بين مجموعة من المتغيرات ، ولا يتضمن العامل تحديدا لأي معنى سيكولوجي أو أسباب، ولايقترح فروضا أو يثبتها ، وقد وجد بعض علماء النفس وجهة النظر هذه جدا جذابة.

ويعتقد آخرون فى عكس هذا الرأى ، فيرون أن التحليل العاملي يقترح فروضا، وكلا نجح فى هذه المهمة انتهت وظيفة الوصف ليصبح جزء من النظرية السيكولوجية من حيث هو

الاحصاء يختصر العلاقات بين مجموعة من المتغيرات، ويقترح علاقات سببية لم يسبق اكتشافها، وأن توليد الفروض ليس حكرا على التحليل العاملي، فهو يشبه فى ذلك طرق الملاحظة والعمل الاكلينيكي ، الا أن الاخيرين يقلان عنه فى درجة الدقة والصرامة. وقد يسهل تكوين الفروض فى مجـال تتوفر فيه ملاحظات كثيرة الا أن اسهام التحليل العاملي يصبح مهما جدا فى المجالات الجديـدة نسبيا، وذلك في الإسراع بتكوين فروض معقولة واستبعاد الفروض الضعيفة. ويتصل هـذا الهدف بإثبات الفروض أو دحضها وبخاصة الفروض المتعلقة بتركيب الشخصية وتنظيمها كفـروض الانمـاط والسمات، مما يصعب اثباته أو دحضه بالطرق غير العاملية. وبين مستوى اقتـراح الفـروض والتحقق منها رابطة متينة ، وقد نجد النوعين من العوامل فى دراسة واحدة. وحيث أن التحليل العاملي يهدف إلى تحقيق واحد أو أكثر من هذه الأهداف الهامة والجوهرية والـتى تتسق مع أهداف العلم الأساسية ، فقد أصبح التحليل العاملي منهجا احصائيا له أساس منطقـي لا غنى عنه فى عدد غير قليل من النظريات السيكولوجية وبالتحديد فى مجال الشخصية التى تدعى عاملية (بدر الانصارى ،1997).

5-11 طرق التحليل العاملي :

تحـدد الطـرق الحسابية المستخدمة فى التحليل العاملي كثيرا، فهناك الطريقة القطرية، والطـريقة المركـزية، والطـريقة المركزية بإستخدام متوسط الارتباطات، وطريقة المكونات الاساسية، ونوجزها فيما يلي:

Factor Analysis: Extraction

Method: Principal components

Analyze – Correl… Covari…

Principal components
Unweighted least squares
Generalized least squares
Maximum likelihood
Principal axis factoring
Alpha factoring

:

1- الطريقة القطرية : Diagonal Method وتعد الطريقة القطرية من الطرق المباشرة والسهلة فى التحليل العاملي، ويمكن استخدامها إذا كان لدينا عدد قليل من المتغيرات وتؤدى إلى اسـتخلاص أكبر عدد ممكن من العوامل وتتطلب هذه الطريقة معرفة سابقة ودقيقة بقيم شيوع

المـــتغيرات ، وبدون هذه المعرفة لايمكن استخدامها، وتستمد الطريقة القطرية اسمها من كونها تقـــوم علـــى اســـتخدام القيم القطرية فى المصفوفة الارتباطية مباشرة . وتبدأ الطريقة القطرية باســـتخلاص هـــذه القيمة بكاملها فى العامل الأول، وبذلك يكون جذر هذه القيمة هو تشبع المتغير الأول على العامل الأول، ويطلق عليه اسم التشبع القطري وهكذا.

2- الطـــريقة المركزية : Centroid Method كانت الطريقة المركزية " لثرستون " أكثـــر طرق التحليل العاملي استخداما وشيوعا إلى عهد قريب نظرا لسهولة حسابها فضلا عن استخلاص عدد قليل من العوامل العامة . غير أن هذه الطريقة تفتقر إلى عدد من المزايا الهامة ، أهمهـــا أنها لا تستخلص الا قدرا محدودا من التباين الارتباطي ، تتحدد قيم الشيوع فى المصفوفة الإرتباطـــية وفـــق تقديـــرات غير دقيقة حيث تستخدم أقصى ارتباط بين المتغير وأى متغير فى المصفوفة وهو اجراء يؤدى إلى خفض رتبة المصفوفة .

3- الطريقة المركزية باستخدام متوسط الارتباطات : Averoid Method لا تختلف هـــذه الطريقة عن الطريقة المركزية المعتادة إلا فى استخدامها تقدير الشيوع عبارة عن متوسط ارتـــباطات المـــتغير ببقـــية المتغيرات في المصفوفة ثم حساب العوامل بعد وضع المتوسط الخاص بارتـــباطات كـــل متغير فى خليته القطرية ولهذا السبب يطلق على هذا الاسلوب اسم الطريقة المركـــزية بإستخدام المتوسطات. غير أن هذه الطريقة لا توفر نفس الدقة التى تجدها فى الطريقة المركزية التامة ، إذ تؤدي إلى خفض محدود فى نسبة التباين التى تعبر عنها العوامل الناتجة. غير أن هـــذه الطريقة تبدو مفيدة فى حالة وجود عدد كبير من المتغيرات دون توفر وسائل آلية لاجراء العمليات الحسابية.

4- أســـلوب تحلـــيل المركبات الأساسية Principal Components Analysis (P.C.A.)

طـــريقة المكـــونات الأساســـية : Principal Componants تعد طريقة المكونات الأساسية التى وضعها "هويتلنج Hottelling "عام 1933 من أكثر طرق التحليل العاملي دقة وشـــيوعاً في بحـــوث الشخصية ، ولهذه الطريقة مزايا عدة منها أنها تؤدي إلى تشبعات دقيقة. وكـــذلك " فإن كل عامل يستخرج أقصى كمية من التباين (أي أن مجموع مربعات تشبعات العامل تصل إلى أقصى درجة بالنسبة لكل عامل) ، وتؤدى إلى أقل قدر ممكن من البواقي، كما

أن المـــصفوفة الارتباطية تختزل إلى أقل عدد من العوامل المتعامدة (غير المرتبطة) و لم تلق طريقة المكـــونات الأساسية فى البداية قبولا كبيرا بين الباحثين نظرا لحاجتها إلى وقت حسابات طويل لإتمامهـــا ولـــذا كان من المستحيل استخدامها يدويا فى حالة المصفوفات الكبيرة ، ولكن بعد الاعـــتماد على الآلات الحاسبة الالكترونية ذات السرعة الفائقة والدقة الشديدة وطاقة التخزين الكبيرة ، اصبحت هذه الطريقة الآن من بين أكثر الطرق شيوعا نظرا لدقة نتائجها بالمقارنة ببقية الطرق.

يعـــتمد اسلوب تحليل المركبات الأساسية (P.C.A.) بصفة اساسية على تفسير وتحليل مجمـــوعة التغايـــرات والتباينات بين البيانات من خلال مجموعة صغيرة من التوليفات الخطية في المـــتغيرات الأساسية. ومن ثم فإن الهدف الأساسى لهذا الأسلوب التحليلى هو تفسير البيانات ومعـــرفة مدى اختلافها وأسباب هذا الاختلاف، وكذلك التعامل مع البيانات بصورة مختصرة Data Reduction مـــن خلال أقل عدد ممكن من العلاقات الخطية والتي تفسر في مجملها أكبر جزء ممكن من الاختلافات والتباينات بينها.

بفرض أن لدينا أ من المتغيرات $\underline{X} = (X_{1j}, ..., X_{pj}, j = 1, ..., N)$ لمجتمع حجمه N من المشاهدات لـــ أ من المتغيرات في الصورة التالية:

$$X = \begin{bmatrix} X_{11} & X_{12} & ..... & X_{1N} \\ X_{21} & X_{21} & ..... & X_{2N} \\ : & : & ..... & : \\ X_{n1} & X_{n2} & ..... & X_{pN} \end{bmatrix}$$

حـــيث $n > p$ وإذا كانت مصفوفة التباينات والتغايرات هى S ، فإن المشكلة تكمن في إيجاد العلاقات أو التوليفات الخطية[1]

(1) Sturt, M., (1982), "A Geometric Approach to Principal Components Analysis". The American Statistician, 36, 365 - 367.

$$
\begin{aligned}
y_1 &= a_{11}X_1 + a_{21}X_2 + \ldots\ldots + a_{p1}X_p = \underline{a}_1' \underline{X} \\
y_2 &= a_{21}X_1 + a_{22}X_2 + \ldots\ldots + a_{p1}X_{p2} = \underline{a}_2' \underline{X} \\
&\vdots \\
y_p &= a_{1p}X_1 + a_{2p}X_2 + \ldots\ldots + a_{pp}X_p = \underline{a}_p' \underline{X}
\end{aligned} \qquad (3\text{-}1)
$$

حيث

$$\text{var}(y_1) = a_1' \Sigma a_1 \qquad 1 = 1, 2, \ldots, p \qquad (3\text{-}2)$$

$$\text{cov}(y_1, y_k) = a_1' \Sigma a_k \qquad 1, k = 1, 2, \ldots, p \qquad (3\text{-}3)$$

وتكـــون المركبات الأساسية هى تلك التوليفات الخطية غير المرتبطة (Uncorrelated linear combination) $y_1, y_2, \ldots\ldots, y_p$ التي تجعل التباينات (3-2) أكبر ما يمكن.

ومـــن ثم يكون المركب الأساسى الأول هو التوليفة الخطية ذات أعلى تباين أى التوليفة الخطية التي تحقق القيمة العظمى للتباين $\text{var}(y_1) = a_1' \Sigma a_1$ بشرط ان $a_1' a_1 = 1$.

ويكـــون المـــركب الأساســـى الثانى هو التوليفة الخطية ذات أعلى التباينات المتبقية في البيانات، أى التوليفة الخطية التي تحقق القيمة العظمى للتباين.

$$\text{var}(y_2) = a_2' \Sigma a_2 \qquad (3\text{-}4)$$

تحت شروط

$$a_2' a = 1$$

$$\text{cov}(a_1' X, a_2' X) = 0$$

وهكـــذا يكون المركب الأساسى رقم 1 هو التوليفة الخطية $a_1' X$ التي تحقق القيمة العظمى للتباينات المتبقية أى العلاقة الخطية $Y_1 = a_1' X$ التي تحقق القيمة العظمى للتباين .

$$\text{var}(y_1) = a_1' \Sigma a_1 \qquad (3\text{-}5)$$

تحت شروط

$$a_1' a_1 = 1$$

$$\text{cov}(a_1' X, a_K' X) = 0 \qquad \text{for} \qquad k < 1$$

وهكذا حتى نحصل على عدد من التوليفات يستحوذ على أكبر قدر مكن من التبيانات.

ويمكـــن التوصل إلى هذه العلاقات باستخدام فكرة مضروب لاجرانج حيث أنه بالنسبة للمركب الاساسى الأول فإن :

$$\begin{aligned} \phi_1 &= \operatorname{var}(y_1) - \lambda_1(a_1' a_1 - 1) \\ &= a_1' \Sigma a_1 - \lambda_1(a_1' a_1 - 1) \end{aligned} \tag{3-6}$$

حـــيث l_1 هـــى مضروب لاجرانج للمركب الأساسى الأول . وبإجراء التفاضلات الجزئية فإن :

$$\begin{aligned} \frac{\partial \phi_1}{\partial a_1} &= 2\Sigma a_1 - 2\lambda_1 a_1 = 0 \\ \frac{\partial \phi_1}{\partial \lambda_1} &= a_1' a_1 - 1 = 0 \end{aligned} \tag{3-7}$$

ومن ثم نحصل على الآتى :

$$(S - l_1 I)\, a_1 = 0 \tag{3-8}$$

وللوصول إلى حل العلاقة (3-8) بشرط ان $a_1' a_{1_1} = 1$ من

$$|S - l_1 I| = 0 \tag{3-9}$$

ومن ثم يمكن التوصل إلى كل من الجذر الكامن الأول (eigenvalue) l_1 والمتجه الكامن الأول (eigenvector) a_1 والنماظر للجذر l_1 .

أما بالنسبة للمركب الأساسى الثانى $y_2 = a_{2_1}' X$ فإن

$$\begin{aligned} \phi_2 &= \operatorname{var}(y_2) - \lambda_2(a_2' - 1) - \mu(a_2' a_2)_1 \\ &= a_2' \Sigma a_2 - \lambda_2(a_2' a_2 - 1) - \mu(a_2' a_2) \end{aligned} \tag{3-10}$$

حـــيث l_2 هو مضروب لاجرانج للمركب الأساسى الثانى وبإجراء التفاضلات الجزئية فإن:

$$\frac{\partial \phi_2}{\partial a_2} = 2\sum a_2 - 2\lambda_2 a_2 - \mu a_1 = 0$$

$$\frac{\partial \phi_2}{\partial \lambda_2} = a_2' a_2 - 1 = 0 \qquad (3\text{-}11)$$

$$\frac{\partial \phi_2}{\partial \mu} = a_1' a_2 = 0$$

ومن ثم نحصل على العلاقة التالية :

$$(S - l_2 I)\, a_2 = 0 \qquad (3\text{-}12)$$

وللوصول إلى حل العلاقة (3-12) بشرط $a_1' a_2 = 0$, $a_2' a_2 = 1$ من

$$|S - l_2 I| = 0 \qquad (3\text{-}13)$$

ومن ثم يمكن التوصل إلى كل من الجذر الكامن الثانى (eigenvalue) l_2 والمتجه الكامن الثانى (eigenvector) a_2 والمناظر للجذر l_2 .

وهكذا يمكن الحصول على المركب الأساسى رقم l ، كما أن a_1 هو المتجه الكامن المناظر له، ومن ثم يكون لدينا عدد من الجذور الكامنة بحيث أن :

$$l_2 \geq l_2 \geq l_3 \geq \ldots\ldots \geq l_p$$

ويناظـــر كل جذر منها متجه كامنا يمثل ثوابت العلاقة الخطية الجديدة، مع ملاحظة أنه علـــى الـــرغم مـــن أنه يكون لدينا عدد من التوليفات الخطية مساويا p ، فإن معظم التباينات والاختلافات بين البيانات والمشاهدات ترجع إلى عدد أقل من التوليفات وليكن m توليفة (m مركب اساسى) حيث أن معظم المعلومات يرجع تفسيرها إلى هذا العدد الصغير من المركبات، ومـــن ثم يمكـــن الاستعاضة عن p بمتغير بعدد m مركبة أساسية لتفسير الظاهرة محل الدراسة، وبالتالى يمكن التعامل مع مجموعة من المعلومات والبيانات حجمها n x m بدلا من n x p .

والجديـــر بالذكـــر أن من صفات أسلوب تحليل المركبات الاساسية صفة التغير والتنوع للعلاقات الخطية الناتجة إلا أنه يختص بخاصيتين وهما[1] :

---

[1] Jackson, J.E., (1980): Principal components and factor analysis: Part I - Principal Components", Journal of Quality Technology, 12, 201-213.

$$(1)\quad |S| = |l| \qquad\qquad (3\text{-}15)$$
$$(2)\quad tr\,S = tr\,l$$

حيث :

مجموع التباينات $tr\,S = s_{11} + s_{22} + ... + s_{pp}$

مجموع الجذور الكامنة $tr\,l = l_1 + l_2 + ... + l_p$

أي أن مجمـــوع التباينات تساوى مجموع الجذور الكامنة eigenvalues ، كما أن محدد مصفوفة التباينات تساوى محدد الجذور الكامنة.

وبالتالى يمكن قياس درجة الأهمية النسبية لكل من مركب اساسى لتفسير سلوك المتغيرات الأصلية في تشكيل الظاهرة باقل خطأ ممكن للمعلومات في النظام كالتالى :

$$RI_i = \frac{\lambda_i}{tr\Sigma} \qquad (3\text{-}16)$$

أى أن الأهمية النسبية (Ri_i) للمركبة الأساسية I مثلا = الجذر رقم I الكامن / مجموع المتباينات

حـــيث تمثل الأهمية النسبية هنا إجمالى التباينات بين المشاهدات المفسرة باستخدام المركبة الأساســية رقـــم I يلاحـــظ أنه في أغلب الأحيان أن حوالى من 80% إلى 90% من إجمالى التبايـــنات لعـــدد كبير من المتغيرات p يمكن رجوعه إلى المركبة الأساسية الأولى أو المركبتين الأساسيتين الأوليين أو المركبات الثلاث الأولى ، ومن ثم يمكن الاستعاضة عن العدد الأصلى p بهذه المركبات الثلاث دون فقد يذكر للمعلومات المتاحة.

طريقة المكونات الاساسية Principal components

يـــشير صفوت فرج (1980) ان طريق المكونات الاساسية ل هوتلنج من اكثر طرق التحلـــيل العاملى دقة ومميزات، غير ان الكثيرين من الباحثين كانوا يحجمون عن استخدامها لما تتطلـــبة مـــن اجـــراءات طويلة وعمليات حسابية متعددة الا انة ازاء التقدم العلمى الراهن في استخدام الحاسبات الالكترونية الحديثة والفائقة السرعة مثال في البحوث النفسية اصبح من غير

المـــستطاع مقاومـــة اغـــراء استخدام هذه الطرق الدقيقة SAS , SPSS ، وطريقة المكونات الاساســـية لا تفترض تسلسل التباين النوعى فى شكل عوامل نوعية ويدمج هذا التباين فى هذة الطـــريقة فى التباين العام مكونا فئات تصنيفية كبرى تتضمن نسبة ضئيلة من هذا التباين النوعى لا تظهـــر واضـــحة فى العوامل المبكرة الاستخلاص عامليا والتى تعد ذات اهمية كبيرة فى هذا الاسلوب.

يـــضاف الى ذلك ميزة رئيسية فى المكونات الاساسية هى ان كل عامل فيها يستخلص اقصى تباين ممكن ، بمعنى ان مجموع المربعات يصل الى اقصى حدودة فى كل عامل وعلى ذلك تتلخص المصفوفة الارتباطية فى اقلا عدد من العوامل المتعامدة.

وهذا معناة ان اسلوب المكونات الاساسية يتميز بقدرتة على الوصول الى حل يتفق مع محك أو فى مربعات للمصفوفة الارتباطية وهو احد المحكات الرياضية التى تلاقى قبولا وضحا فى مجـــال الاساليب التلخيصية للعلاقات بين المتغيرات تهتم بعملية شرح وتفسير بناء التباين (فرج، 1980) . وطـــريقة المكونات الاساسية Principal components وبناء التباين التلازمى أو التغاير المزدوج Variance structure

مـــن خـــلال الارتـــباطات الخطية القليلة Covariance structure بالمتغيرات الاصلية وموضوعاتة العامة هي :Linear combination اختزال او تحويل البيانات Data reduction الشرح او التفسير Interpretation (Johenson, & Wichern , 1992 : 256)

1. التدوير المتعامد والمائل

هـــناك نوعان من التدوير تبعا للزاوية التى تفصل بين المحاور المرجعية وهما المتعامد والمائل Orthogonal، ففى التدوير المتعامد تدار العوامل معاً (اثنين منهما مثلا) مع الاحتفاظ بالتعامد (90 درجة) اما التدوير المائل oblique ففية تدار المحاور دون احتفاظ بالتعامد ، فنترك لتتخذ الميل الملائم لها .

وفى هذا الصدد يذكر محمود منسى (1993) ان جميع العوامل المستخلصة من التحليل العاملى المباشر ، تحتاج الى ابراز هويتها بطريقة اوضح لانة يصعب تفسيرها سيكولوجيا وحيث ان هدف تدوير المحاور (المتعامد والمائل على حد السواء) هو التوصل الى البناء البسيط الا ان

التدويـــر المـــتعامد يهـــدف الى تحقيق هذا الهدف فى ضوء فكرة الاستقلال بين العوامل وعدم الارتباط (جتا 90= 0) (منسى ، 1993 : 70)

والعوامل المتعامدة غير مرتبطة معا ، اى ان معاملات الارتباط بينها تساوى الصفر ، اذ تصنف العوامل الاختبارات أو المتغيرات الى فئات غير مرتبطة ، وهكذا يصبح التقسيم حاد غير مـــتداخل ، امـــا العوامل المائلة فهى بينها ارتباط اى الها عوامل متداخلة ويفضل بعض المحللين استخراج عوامل متعامدة غير مرتبطة فى حين يهتم آخرون باستخلاص المائل ، ويهدف تدوير المحاور الى تحقيق ما يسمية ثرستون البناء البسيط (عبد الخالق ، 1987 : 116).

بعض طرق التدوير المتعامد

- الكوارتيماكس
- الباريماكس
- الماكسبلان

بعض طرق التدوير المائل

- الكوارتيمن
- الاوبلمن
- الكوفاريمين
- البروماكس

محكات التوقف عن استخلاص العوامل : (تدوير العوامل)

1- **محك تايكر Tuker Phi** : وهو يقوم على استخدام معامل فاي ويعتمد على مبدأ انـــه اذا لم يوجد نقص ذي دلالة في حجم قيم البواقي من مصفوفة الى اخرى تليها فان العامل الـــذي استخلص يكون ذا دلالة وما يتبقى ليس الا بواقي لا اهمية لها. ويلاحظ ان هذه الطريقة تـــصلح بالنسبة للطريقة المركزية لثيرستون ونعتمد صحة هذا المحك جزئيا على صحة اجراءات عكـــس الاشـــارات في المصفوفة كما ان عملية العكس تؤدي الى ارتفاع في القيم الايجابية مما يترتب عليه ان تتجاوز قيمة فاي قيمة (ن-1)/ (ن+1) .

2- **قاعـــدة همفـــري** : تعـــتمد هـــذه القاعدة على حجم العينة الاصلية التي حسبت الارتـــباطات بين متغيراتها وثانيا على فكرة ان تشبع متغيرين فقط دون المصفوفة كلها كافيين تماما لتقرير وجود عامل عام .

3- **محك كومب** : يطبق هذا المحك فقط على المصفوفات التي تحتوي على قيم موجبة او صـــفرية ويسمح بالقيم السالبة الصغيرة التي لا تختلف اختلافا واضحا عن الصفر وبذلك يعتمد هذا الاسلوب على نمط البواقي في المصفوفة اكثر من اعتماده على حجمها او دلالاتها حيث انه يفتـــرض انه في حالة وجود عوامل ذات دلالة مرتفعة لم تستخلص بعد وليس مجرد تباين خطأ المـــصفوفة فعليـــنا ان لا نتوقع قيم سالبة اكثر في مصفوفة البواقي بعد العكس مما يتوقع بحكم الصدفة في مصفوفة ناتجة عن ارتباطات ايجابية .

4- **محـــك كايـــزر** : يعتمد هذا المحك على حجم التباين الذي يعبر عنه العامل ، وعلى ذلك فان هذا المحك يتطلب مراجعة الجذر الكامن للعوامل الناتجة ، وعلى ان تقبل العوامل التي يـــزيد جـــذرها الكامن عن الواحد الصحيح ، وتعد عوامل عامة . وهو محك قد يكون صالحا ومناسبا لطريقة المكونات الاساسية لهوتلنج على وجه الخصوص . ويذكر عبد الخالق (1994) ان العـــوامل الدالة في هذه الطريقة هي العوامل التي يساوي او يزيد جذرها الكامن على واحد صحيح أي ان التباين الذي يستوعبه كل عامل (مجموع مربعات التشبعات على كل عامل) $\leq$ 1.. بشرط ان يكون قد وضع في الخلايا القطرية واحد صحيح ويذكر وايت وزملاؤه ان هذا المعيار تتطابق نتائجه مع معايير اخرى .

5- **محـــك كاتل** : تؤدي خطوات استخلاص العوامل من المصفوفة الارتباطية الى انتاج العـــوامل الاكثر عمومية اولا في كل الاساليب العاملية بلا استثناء ، ثم تبدا العوامل الخاصة او التـــباين النوعي في الظهور. وفي طريقة كالمكونات الاساسية لاتفرق بين عوامل عامة واخرى غـــير عامـــة يفتـــرض ايضا ان حجم التباين النوعي الذي يتسرب الى العوامل الناتجة يتزايد في العـــوامل الاخيرة ويبدأ في فرض صورةتقلل من اهمية المصفوفة العاملية ويتطلب الامر في هذه الحالة تحديد العدد الامثل من العوامل قبل ان تؤدي ظهور التباينات الخاصة الى احداث خلل في مصفوفة العوامل ويقترح كاتل هنا محكا بسيطا يطلق عليه اسم البقايا المبعثرة وذلك بان نقوم برسم محورين متعامدين ، افقي نضع عليه عدد العوامل في تحليلنا (الذي انتج فيه عددا كبيرا من

العـــوامل) ويقـــسم المحـــور الراسي وفقا لوحدات منتظمة معبرة عن الجذر الكامن المستخلص للعوامل المختلفة.

وســـنلاحظ بعد اتمام رصد عواملنا وجذورها الكامنة ، ان حجم الجذر يتناقض بشكل كبير في العوامل الاولى الى ان يصل الى نقطة معينة هي غالبا حول جذر كامن واحد صحيح ثم يبدا حجم الجذر في التناقص بصورة ضئيلة بحيث يستوي فيها الخط البياني مع الخط الافقي .

وإذا افترضنا أن النقطة التي سنتوقف لديها في قبولنا للعوامل هي عند العامل الرابع على ســـبيل المثال فان الفرق لن يكون كبيرا في الواقع بين ما يقدمه محك كاتل وبين ما يقدمه محك كايزر الذي يتطلب التوقف عند العامل الثالث هذا على سبيل المثال .

وتتبقى لطريقة كايزر ميزتها في هذه الحالة في كونها لا تتطلب استخلاص عدد كبير من العـــوامل ثم رصدها في الشكل البياني للتعرف على نقطة توقف التناقض واستواء الخط ، حيث يمكـــن حـــساب الجذر الكامن لكل عامل بطريقة كايزر قبل استخلاص العامل التالي مما يوفر جهدا لا مبرر له (صفوت فرج ، 1991 : 246) .

6- **محك مويزز** : يقوم هذا المحك على تفرطح التباين الكلي للعوامل المتتالية .

7- **محـــك بـــيرت وبانكـــز** : ويمكن عن طريق هذا المحك تحديد العوامل ذات الدلالة المنخفـــضة عن طريق تحديد الخطأ المعياري للتشبع الصفري ، وبمقارنة عدد تشبعات العامل أو مضاعفات هذا العدد التي يزيد مقدارها عن الخطأ المعياري .

6-11 بعض مفاهيم التحليل العاملي:

1- درجة الشيوع Communality:

تعـــرف درجـــة شـــيوع المتغير بإسهامات هذا المتغير في جميع العوامل ويقاس بمجموع مـــربعات معاملات هذا المتغير في العوامل المختلفة، فمثلا تقاس درجة شيوع المتغير رقم (j) على النحو التالي:

$$C_j = \sum_{i=1}^{m} \alpha_{ij}^2$$

درجة التشبع Loading :

α_{ij}

يعرف المعامل

بمعامـــل تحمـــيل (أو تـــشبع) المتغير i على العامل j كما يعبر عن مدى ارتباط العامل بالمتغير. ويلاحظ أن مجموع مربعات درجات التشبع لكل عامل تسمى الجذر الكامن وتعبر عن أهمية هذا العامل في تفسير الاختلافات في المتغيرات، كما يعبر مجموع الجذور الكامنة عن التباين الـــذي أمكـــن تفسيره من خلال العوامل، وبنسبته إلى عدد المتغيرات نحصل على نسبة التباين العاملـــية هـــذه ؛ لأنـــه يتم تحويل المشاهدات إلى قيم معيارية ويكون تباين كل متغير الواحد الصحيح .

2- معامل الارتباط كعلاقة بين الاختبار (1) والاختبار(2) يساوي=

حاصـــل ضرب درجة تشبع الاختبار (1) بعامل معين (أ) × درجة تشبع الاختبار (2) بعامل معين (أ) .

أي أن $ر_{1.2}$= $ش^{1}_{أ}$ × $ش^{2}_{أ}$

وقياسا على ذلك فان معامل الارتباط بين الاختبار(1) ونفسه =$(ش^{1}_{أ})^{2}$

3- التباين Variance:

يحسن الاعتماد في التحليل العاملي على الدرجات المعيارية z- Score وهي تعني توحيد أســـاس الدرجات على المتغيرات المختلفة بحيث تصبح وحدة الدرجة الخاصة بالفرد على المتغير واحد صحيح أو درجة أي فرد عبارة عن نسبة من هذا الواحد الصحيح وبهذا نلاحظ ان تباين المتغير الواحد في أي تحليل عاملي هو:

التباين الكلي = تباين العامل العام + تباين العامل النوعي + تباين الخطأ

حـــيث أن: تباين المتغير العام هو: مربع تشبع المتغير أو مربع **ارتباطه بالعامل وهو هنا تباين** عام يشترك به المتغير مع تباينات لمتغيرات أخرى.بما يؤدي إلى استخلاص عامل عام

تباين المتغير النوعي هو: مربع تشبع المتغير أو مربع ارتباطه بالعامل وهو قدر من التباين الذي يعبر به المتغير الواحد عن نوعية أدائه ويظهر على عامل دون أن يظهر معه تباين لمتغيرات أخرى .

تباين الخطأ هو :الجزء الذي لا يستخلص في شكل عوامل ويبقى في المصفوفة الارتباطية بعد استخلاص العوامل على شكل بقايا ويعود إلى عدد من الأسباب وهي:

أ- **أخطـــاء القـــياس** : ويقصد بها استخدام الأدوات منخفضة الثبات أو استخدام مقاييس غير متجانسة البنود ، أو تأثير بعض المتغيرات الأخرى فكل هذا يؤثر على نتائج التحليل العاملي .

ب- **أخطاء التجربة** : والتي تتمثل في عدم الضبط الدقيق للمتغيرات بالبحث .

جـــ - **أخطـــاء الدقة** : والتي تتمثل في عدم إحكام جلسة الاختبار أو طريقة تقديم التعليم أو أسلوب تصحيح الاختبارات .

4- العلاقة بين الثبات والشيوع :

أن معامل الثبات يعبر عن الحجم الحقيقي لتباين المتغير أي بعد استبعاد تباين الخطأ وأننا ننظر إلى قيم الشيوع للمتغير في مصفوفة عاملية باعتبارها معامل ثبات لهذا المتغير حيث تمثل قيم الشيوع في هذه الحالة هذا التباين الحقيقي الذي استخلص معبرا عن تباينات مختلفة يشترك فيها المتغير مع غيره من المتغيرات طالما بقي تباين الخطأ في مصفوفة البواقي معبرا بدوره عن الجزء من التباين الكلي الذي لا يشترك فيه الاختبار مع غيره من المتغيرات نتيجة لأخطاء القياس أو أخطاء التجريب.

5- الجذر الكامن: Eigen Value

يعـــرف مجمـــوع مربعات تشبعات كل المتغيرات على كل عامل على حدة من عوامل المـــصفوفة باسم الجذر الكامن للعامل وهو تعبير يستخدم في جبر المصفوفات ويلاحظ بالنسبة لأي مـــصفوفة عاملية أن الجذر الكامن يتناقص تدريجيا عن العامل الأخر، فالعوامل الأولى ذات جـــذر كامن اكبر من العوامل المتأخرة الاستخلاص ، ذلك أن خطوات حساب العوامل تؤدي إلى استخلاص أقصى تباين مشترك بين المتغيرات في كل مرة على التوالي وبطرح مصفوفة الناتج من المصفوفة الارتباطية يتبقى حجم اصغر من التباين المشترك بين المتغيرات يستخلص في عامل جديد ذي جذر كامن اصغر من سابقه. وسيكون مجموع قيم الشيوع للمتغيرات يساوي تماما

مجمـــوع الجذور الكامنة لعوامل المصفوفة ، بمعنى أخر أن مجموع مربعات الصفوف ((أي قيم الشيوع)) = مجموع مربعات الأعمدة ((أي الجذور الكامنة)).

ولمـــا كـــان التباين الذي يساهم به متغير بقيمه المعيارية يساوي (1) فان أي مكون جذره الكامن اقل من واحد لا يكون له أي أهمية تذكر ، ومعنى ذلك أن المكونات أو العوامل التي تكون قيمة الجذر الكامن لكل منها واحد أو اكثر هي التي تعتمد وتعتبر ذات دلالة معنوية

6- حجم التباين العاملي ونسبة التباين العاملي :

حجم التباين العاملي هو مجموع قيم الشيوع أو مجموع الجذور الكامنة ، أما نسبة التباين العاملي للمصفوفة عبارة عن :

$$\frac{\text{مجموع الجذور الكامنة للعوامل} \times 1}{\text{التباين الارتباطي}}$$

والتباين الارتباطي يساوي عدد المتغيرات التي تدخل في التحليل العاملي والجذر الكامن يعكس مقدار التباين العام عن طريق العدد النسبي من العوامل.

11-7 استخدام برمجية SPSS في حل المسائل.

اذا كـــان لديـــنا استبانة تتكون من 23 سؤالاً حول البرنامج الاحصائي SPSS وهي كالتالي:

استبانة حول الاحصاء والبرنامج SPSS	موافق بشدة	موافق	محايد	معارض	معارض بشدة
1. الاحصاء يجعلني أبكي.					
2. يعتقدأصـــدقائي أنـــي غـــبي لأنني لم ابدأ بالتعامل مع البرنامج الإحصائي للعلوم الاجتماعية.					
3. الانحرافات المعيارية تثيرني.					
4. أحلم بأن بيرسون يهاجمني في معاملات الارتباط.					
5. انا لا أفهم الإحصاء.					
6. ليس لدي الكثير من الخبرة في الحاسوب.					
7. جميع الحواسيب تكن الكراهية لي.					
8. أنا لست جيداً في الرياضيات.					

9. أصدقائي في الإحصاء أفضل مني.					
10. الكمبيوتر مفيدة فقط في اللعب.					
11. فعلت سيئا في الرياضيات في المدرسة.					
12. يحاول الناس القول بأن SPSS يجعل الاحصاءأيسر على الفهم ولكنه بالنسبة لي ليس كذلك.					
13. بسبب استخدام الكمبيوتر إنني أخشى أن يسبب ضررا لا يمكن إصلاحه.					
14. الكمبيوتر يحتاج الى عقول وأنا اشعر بأني على غير ما يرام عند استخدامها.					
15. اشعر بأن الكمبيوتر يرفضني.					
16. أنا أبكي صراحة عند ذكر مقاييس النزعة المركزية.					
17. أنا تصيبني حالة غيبوبة كلما أرى المعادلات.					
18. تتحطم برامج الإحصائي للعلوم الاجتماعية دائما عندما أحاول استخدامها.					
19. الجميع ينظر لي عندما أقوم باستخدام البرنامج الإحصائي للعلوم الاجتماعية.					
20. لا يمكنني النوم عندما افكر في المتجهات الكامنة.					
21. انا اشعر بالسوء عندما افكر بالتوزيع الطبيعي.					
22. أصدقائي في البرنامج الإحصائي للعلوم الاجتماعية على نحو أفضل مما أنا عليه.					
23. إذا كنت جيدا في الإحصاء فإن أصدقائي سوف يفكرون انني أنا الطالب الذي يذاكر كثيرا.					

1. توصيف المتغيرات:

	Name	Type	Width	Decimal	Label	Values	Missing	Columns	Align	Measure
1	q01	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
2	q02	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
3	q03	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
4	q04	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
5	q05	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
6	q06	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
7	q07	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
8	q08	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
9	q09	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
10	q10	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
11	q11	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
12	q12	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
13	q13	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
14	q14	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
15	q15	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
16	q16	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
17	q17	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
18	q18	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
19	q19	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
20	q20	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
21	q21	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
22	q22	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal

2. إدخال البيانات :

	q01	q02	q03	q04	q05	q06	q07	q08	q09	q10	q11	q12	q13	q14	q15	q16	q17	q18	q19	q20	q21	q22	q23
1	2	1	4	2	2	2	3	1	1	2	1	2	2	2	2	3	1	2	3	2	2	2	5
2	1	1	4	3	2	2	2	2	5	2	2	3	1	3	4	3	2	2	3	4	4	4	2
3	2	3	2	2	4	1	2	2	2	2	3	3	2	4	2	3	2	3	1	4	3	2	2
4	3	1	1	4	3	3	4	2	2	4	2	2	2	3	3	3	2	4	2	4	4	4	3
5	2	1	3	2	2	3	3	2	4	2	2	3	3	2	2	2	2	3	3	4	2	4	4
6	2	1	3	2	4	4	4	2	4	3	2	4	3	3	5	2	3	5	1	5	3	1	4
7	2	3	3	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	3	2	2	4	4
8	2	2	3	2	2	2	2	2	4	2	2	3	2	2	3	2	2	2	4	3	2	4	4
9	3	3	1	4	5	3	5	5	3	3	5	5	5	5	5	5	5	5	2	5	5	3	3
10	2	4	4	3	2	1	2	2	3	2	2	3	2	1	2	3	2	2	3	3	2	4	4
11	2	1	5	2	2	1	2	2	5	2	1	3	1	2	1	2	2	2	5	3	2	5	5
12	2	1	3	3	4	3	3	1	3	2	2	3	2	2	3	3	2	2	3	4	3	4	4
13	3	1	3	4	3	2	3	3	2	3	3	4	4	4	4	4	3	3	2	4	4	3	4
14	2	2	1	2	2	3	3	2	2	3	2	4	2	4	4	4	2	4	1	5	5	3	1
15	2	2	3	4	2	2	3	2	2	3	2	3	2	3	3	4	2	3	3	4	4	4	4
16	3	1	2	2	2	2	2	2	2	3	2	3	2	3	2	3	2	3	2	3	2	3	4
17	1	2	5	2	1	1	1	1	4	1	1	2	1	1	1	2	2	1	4	2	1	4	4
18	2	2	3	3	3	4	3	2	5	2	3	3	3	3	4	3	2	2	2	3	3	3	4
19	2	3	4	2	3	1	1	1	5	2	1	3	1	2	2	3	1	1	4	2	2	4	4
20	2	1	1	2	3	4	4	1	5	1	2	5	2	5	5	5	2	5	1	5	5	5	5

3. البدء في التحليل Analyze – Data Reduction - Factor :

saq - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

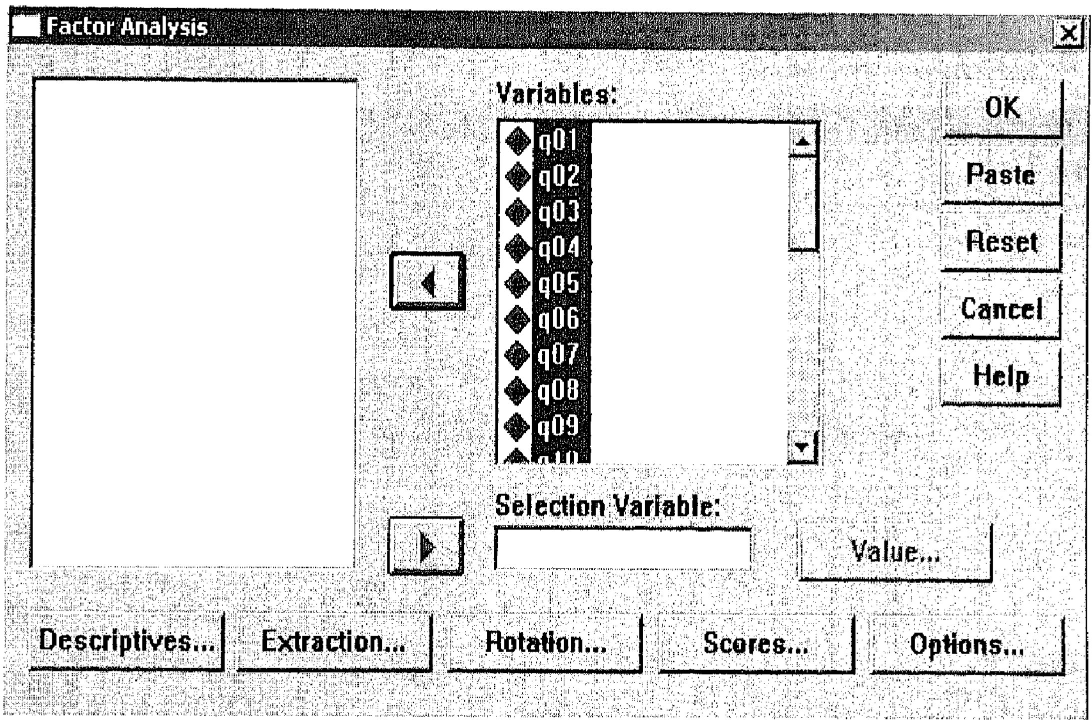
Analyze: Reports ▸ / Descriptive Statistics ▸ / Compare Means ▸ / General Linear Model ▸ / Correlate ▸ / Regression ▸ / Classify ▸ / Data Reduction ▸ Factor... / Scale ▸ / Nonparametric Tests ▸ / Multiple Response ▸

1 : q01

	q01	q02			q05	q06	q07
1	2	1			2	2	3
2	1	1			2	2	2
3	2	3				1	2
4	3	1			3	3	4
5	2	1			2	3	3
6	2	1			4	4	4
7	2	3	3	2	2	2	2
8	2	2	3	2	2	2	2
9	3	3	1	4	5	3	5
10	2	4	4	3	2	1	2

4. إختيار المتغيرات :

نختار المتغيرات ونضعها في قائمة Variables



5. إستكمال محددات التحليل :

نضغط على زر Descriptives – ثم نحدد Initial solution

Factor Analysis: Descriptives

Statistics
- [x] Univariate descriptives
- [x] Initial solution

Continue
Cancel
Help

Correlation Matrix
- [x] Coefficients
- [x] Inverse
- [x] Significance levels
- [x] Reproduced
- [x] Determinant
- [x] Anti-image
- [x] KMO and Bartlett's test of sphericity

– ثم نضغط زر Continue

نضغط على زر Extraction – ثم نحدد:

Factor Analysis: Descriptives

Statistics
- [x] Univariate descriptives
- [x] Initial solution

Continue
Cancel
Help

Correlation Matrix
- [x] Coefficients
- [x] Inverse
- [x] Significance levels
- [x] Reproduced
- [x] Determinant
- [x] Anti-image
- [x] KMO and Bartlett's test of sphericity

– ثم نضغط زر Continue

نضغط على زر Rotation... – ثم نحدد:

Factor Analysis: Rotation

Method
None
Varimax
Direct Oblimin
Delta: 0
Quartimax
Equamax
Promax
Kappa 4

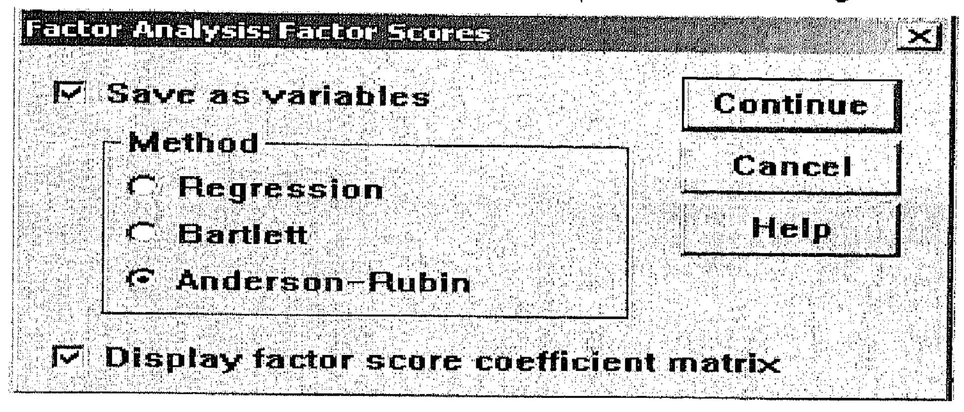
Display
Rotated solution
Loading plot(s)

Maximum Iterations for Convergence: 30

Continue
Cancel
Help

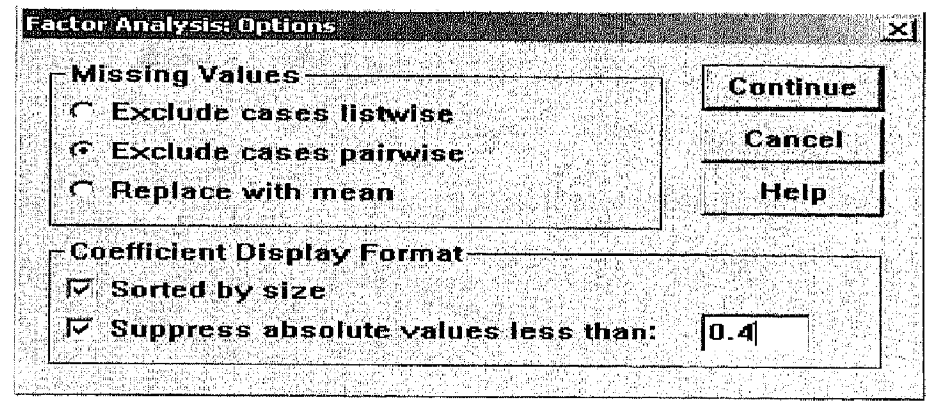
– ثم نضغط زر Continue

نضغط على زر Scores... – ثم نحدد:



– ثم نضغط زر Continue

نضغط على زر Options... – ثم نحدد:



– ثم نضغط زر Continue

6. مخرجات التحليل :

الاحصاءات الوصفية: الوسط الحسابي، والانحراف المعياري للاسئلة.

Descriptive Statistics

	Mean	Std. Deviation	Analysis N	Missing N
Q01	2.37	.828	2571	0
Q02	1.62	.851	2571	0
Q03	2.59	1.075	2571	0
Q04	2.79	.949	2571	0
Q05	2.72	.965	2571	0
Q06	2.23	1.122	2571	0
Q07	2.92	1.102	2571	0
Q08	2.24	.873	2571	0
Q09	2.85	1.263	2571	0
Q10	2.28	.877	2571	0
Q11	2.26	.881	2571	0
Q12	3.16	.916	2571	0
Q13	2.45	.949	2571	0
Q14	2.88	.999	2571	0
Q15	2.77	1.009	2571	0
Q16	2.88	.916	2571	0
Q17	2.47	.884	2571	0
Q18	2.57	1.053	2571	0
Q19	2.29	1.101	2571	0
Q20	3.62	1.036	2571	0
Q21	3.17	.985	2571	0
Q22	2.89	1.041	2571	0
Q23	3.43	1.044	2571	0

مصفوفة الارتباط بين كل ازواج الاسئلة وتستخدم لفحص العلاقة بين الاسئلة، انه عندما يوجد ارتباط كبير نسبيا بين علامات اختبارين فمعنى ذلك أن هناك تباينا مشتركا بين الاختبارين، أي انهما يقيسان شيئا مشتركا، ونلاحظ أن ارتباطاتها بالاختبارات الأخرى متدنية بشكل ملحوظ.

إذن يمكن أن نخلص إلى أن الفكرة الأساسية للنموذج العاملي هي افتراض إمكانية تجميع المتغيرات بناءا على معاملات الارتباط بينها ، هذا يعني أن جميع المتغيرات الموجودة في مجموعة معينة مرتبطة مع بعضها ارتباطا قويا ، ولكن ارتباطها بمتغيرات المجموعات الأخرى ارتباطا

ضـــعيفا ، ومن الممكن أن نتصور أن كل مجموعة من المتغيرات تمثل عاملا واحدا وهو المسئول عن الارتباط المشاهد بينها .

Correlation Matrix

		C01	C02	C03	C04	C05	C06	C07
Correlation	C01	1.000	-.100	-.307	.408	.402	.217	.305
	C02	-.100	1.000	.373	-.172	-.179	-.014	-.159
	C03	-.307	.373	1.000	-.380	-.370	-.227	-.342
	C04	.408	-.172	-.380	1.000	.401	.278	.419
	C05	.402	-.179	-.370	.401	1.000	.257	.319
	C06	.217	-.014	-.227	.278	.257	1.000	.514
	C07	.305	-.159	-.342	.419	.319	.514	1.000
	C08	.311	-.150	-.258	.328	.298	.221	.287
	C09	-.062	.375	.190	-.125	-.098	-.113	-.128
	C10	.274	-.084	-.201	.276	.253	.162	.244
	C11	.367	-.144	-.367	.398	.299	.358	.345
	C12	.345	-.138	-.410	.442	.347	.313	.423
	C13	.366	-.140	-.373	.344	.302	.400	.412
	C14	.338	-.168	-.377	.361	.315	.402	.441
	C15	.291	-.165	-.312	.324	.281	.391	.381
	C16	.408	-.168	-.419	.419	.335	.244	.319
	C17	.311	-.087	-.327	.333	.310	.242	.301
	C18	.347	-.194	-.375	.352	.322	.513	.501
	C19	-.138	.203	.342	-.198	-.165	-.107	-.219
	C20	.274	-.202	-.325	.343	.200	.101	.221
	C21	.338	-.256	-.417	.410	.335	.272	.403
	C22	-.104	.231	.204	-.184	-.120	-.125	-.108
	C23	-.004	.100	.150	-.084	-.072	-.084	-.070
Sig. (1-tailed)	C01		.000	.000	.000	.000	.000	.000
	C02	.000		.000	.000	.000	.000	.000
	C03	.000	.000		.000	.000	.000	.000
	C04	.000	.000	.000		.000	.000	.000
	C05	.000	.000	.000	.000		.000	.000
	C06	.000	.000	.000	.000	.000		.000
	C07	.000	.000	.000	.000	.000	.000	
	C08	.000	.000	.000	.000	.000	.000	.000
	C09	.000	.000	.000	.000	.000	.000	.000
	C10	.000	.000	.000	.000	.000	.000	.000
	C11	.000	.000	.000	.000	.000	.000	.000
	C12	.000	.000	.000	.000	.000	.000	.000
	C13	.000	.000	.000	.000	.000	.000	.000
	C14	.000	.000	.000	.000	.000	.000	.000
	C15	.000	.000	.000	.000	.000	.000	.000
	C16	.000	.000	.000	.000	.000	.000	.000
	C17	.000	.000	.000	.000	.000	.000	.000
	C18	.000	.000	.000	.000	.000	.000	.000
	C19	.000	.000	.000	.000	.000	.000	.000
	C20	.000	.000	.000	.000	.000	.000	.000
	C21	.000	.000	.000	.000	.000	.000	.000
	C22	.000	.000	.000	.000	.000	.000	.000
	C23	.451	.000	.000	.021	.017	.000	.000

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.930
Bartlett's Test of Sphericity	Approx. Chi-Square	19334.492
	df	253
	Sig.	.000

Communalities

	Initial	Extraction
Q01	1.000	.435
Q02	1.000	.414
Q03	1.000	.530
Q04	1.000	.469
Q05	1.000	.343
Q06	1.000	.654
Q07	1.000	.545
Q08	1.000	.739
Q09	1.000	.484
Q10	1.000	.335
Q11	1.000	.690
Q12	1.000	.513
Q13	1.000	.536
Q14	1.000	.488
Q15	1.000	.378
Q16	1.000	.487
Q17	1.000	.683
Q18	1.000	.597
Q19	1.000	.343
Q20	1.000	.484
Q21	1.000	.550
Q22	1.000	.464
Q23	1.000	.412

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.290	31.696	31.696	7.290	31.696	31.696	3.730	16.219	16.219
2	1.739	7.560	39.256	1.739	7.560	39.256	3.340	14.523	30.742
3	1.317	5.725	44.981	1.317	5.725	44.981	2.553	11.099	41.842
4	1.227	5.336	50.317	1.227	5.336	50.317	1.949	8.475	50.317
5	.988	4.295	54.612						
6	.895	3.893	58.504						
7	.806	3.502	62.007						
8	.783	3.404	65.410						
9	.751	3.265	68.676						
10	.717	3.117	71.793						
11	.684	2.972	74.765						
12	.670	2.911	77.676						
13	.612	2.661	80.337						
14	.578	2.512	82.849						
15	.549	2.388	85.236						
16	.523	2.275	87.511						
17	.508	2.210	89.721						
18	.456	1.982	91.704						
19	.424	1.843	93.546						
20	.408	1.773	95.319						
21	.379	1.650	96.969						
22	.364	1.583	98.552						
23	.333	1.448	100.000						

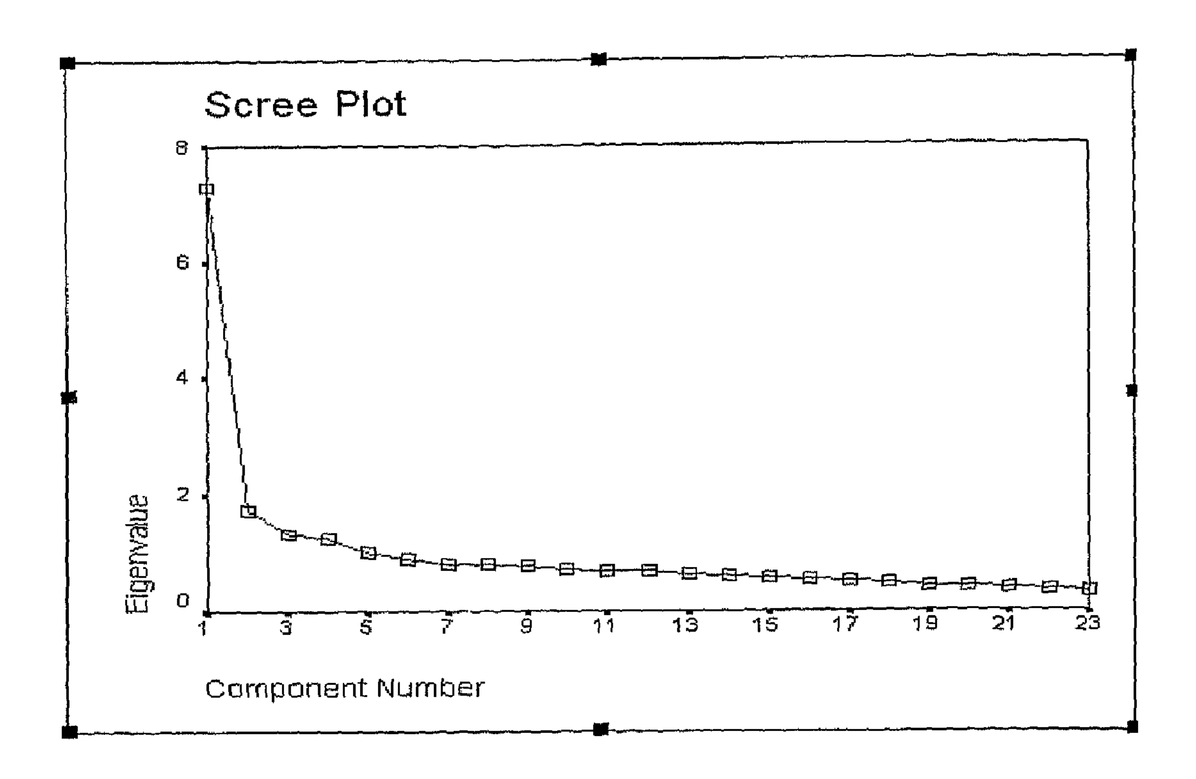
Extraction Method: Principal Component Analysis.

جدول رقم (2)

العوامل المكونة للأداة وجذورها الكامنة ونسبة التباين العاملية

رقم العامل	الجذر الكامن	نسبة التباين العاملية	نسبة التباين التراكمية
1	7.290	31.696	31.696
2	1.739	7.560	39.256
3	1.317	5.725	44.981
4	1.227	5.336	50.317
22	0.364	1.583	98.562
23	0.333	1.448	100.00

وحـــيث تم اجــراء التحليل العاملي كأحد إجراءات تحقيق صدق الأداة لذا يتضح من النتائج الواردة في جدول رقم (2) أن الأداة تتكون من ثلاث وعشرون عاملاً تتشبع عليها بنود الأداة بقـــيم تفـــوق (0.30) حسب محك جيلفورد أما الجذور الكامنة للعوامل فتتراوح بين 7.290 و0.333 بيـــنما كانـــت نسبة التباين العاملية بين 31.696 و1.448 وهذه قيم مقـــبولة حسب محك كيزر. وهذه العوامل مجتمعة تفسر ما نسبته (50.317%) من الظاهرة وهذه النسبة عالية خاصة إذا علمنا أن نسبة (10%) تعتبر نسبة مقبولة.



Component Matrix[a]

	Component			
	1	2	3	4
Q18	.701			
Q07	.686			
Q16	.679			
Q13	.673			
Q12	.669			
Q21	.668			
Q14	.656			
Q11	.652			-.400
Q17	.643			
Q04	.634			
Q03	-.629			
Q15	.593			
Q01	.586			
Q05	.556			
Q08	.549	.401		-.417
Q10	.437			
Q20	.436		-.404	
Q19	-.427			
Q09		.627		
Q02		.548		
Q22		.465		
Q06	.562		.571	
Q23				.507

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

Rotated Component Matrix[a]

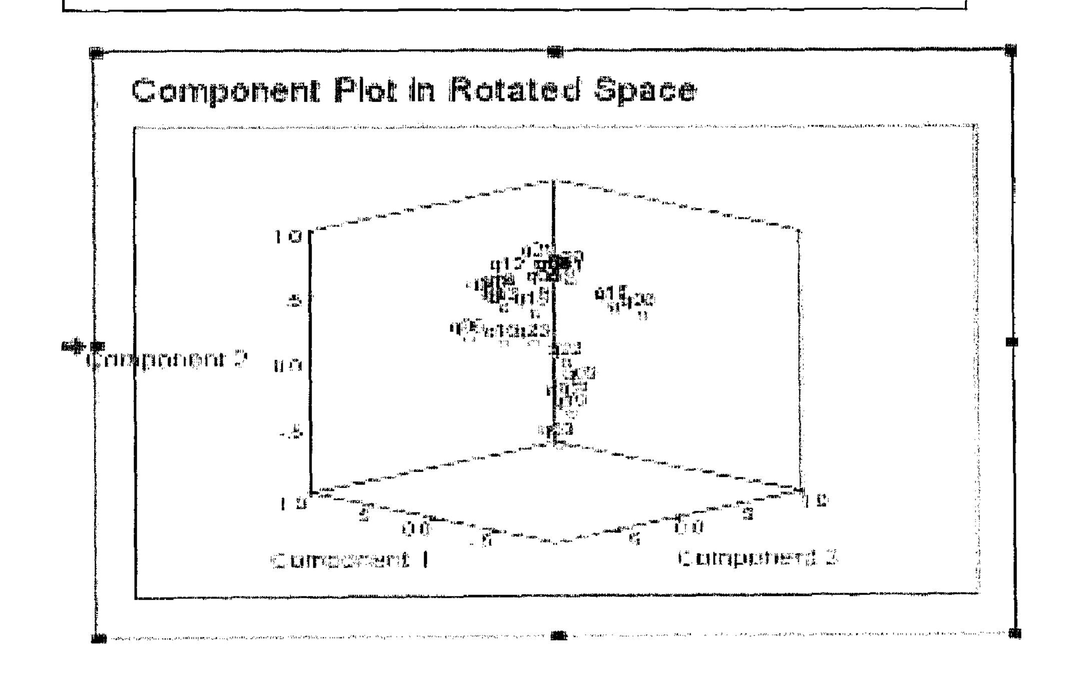
	Component			
	1	2	3	4
Q06	.800			
Q18	.684			
Q13	.647			
Q07	.638			
Q14	.579			
Q10	.550			
Q15	.459			
Q20		.677		
Q21		.661		
Q03		-.567		
Q12	.473	.523		
Q04		.516		
Q16		.514		
Q01		.496		
Q05		.429		
Q08			.833	
Q17			.747	
Q11			.747	
Q09				.648
Q22				.645
Q23				.586
Q02				.543
Q19				.428

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Transformation Matrix

Component	1	2	3	4
1	.635	.585	.443	-.242
2	.137	-.167	.488	.846
3	.758	-.513	-.403	.008
4	.067	.605	-.635	.476

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.



8-11 تمارين Exercise

اذا كان لديك البيانات التالية جد العوامل التي تكونها البيانات.

	stno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1
21	21	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
22	22	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
23	23	2	2	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1
24	24	2	2	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0
25	25	2	3	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1
26	26	2	3	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
27	27	2	3	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1
28	28	2	3	1	1	0	1	0	0	1	0	1	1	1	0	1	1	1
29	29	2	3	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
30	30	2	3	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

الفصل الثاني عشر

الإحصاءات اللامعلمية

NONPARAMETRIC STATISTICS

1-12 مـقـدمـة Introduction

2-12 الطرق اللامعلمية (عينة واحدة)
Nonparametric Methods (Single Sample)

3-12 الطرق اللامعلمية (عينتين مستقلتين)
Nonparametric Methods (Two Independent Samples)

4-12 الطرق اللامعلمية (عينتين مرتبطتين)
Nonparametric Methods (Two Related Samples)

5-12 الطرق اللامعلمية (ثلاثة عينات مستقلة أو أكثر)
Nonparametric Methods (3 or more Independent Samples)

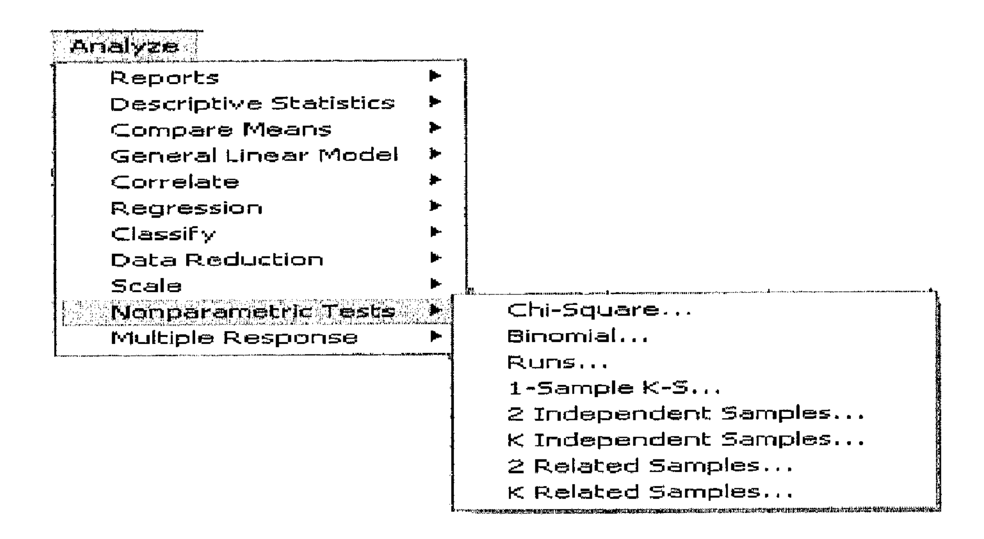
6-12 الطرق اللامعلمية (ثلاثة عينات مرتبطة أو أكثر)
Nonparametric Methods (3 or more Related Samples)

7-12 استخدام برمجية SPSS في حل المسائل.

8-12 تـمـاريـــن Exercise.

الفصل الثاني عشر

الإحصاءات اللامعلمية
NONPARAMETRIC STATISTICS



1-12 مقدمة Introduction

إذا كانت المتغيرات التابعة مقاسة بمقياس اسمي أو رتبي أو عندما لا نتمكن من الإيفاء بافتراضات الاختبارات المعلمية، فأنه يمكن استخدام الاختبارات اللامعلمية Nonparametric tests لأن مثل هذه الاختبارات لا تتطلب أية افتراضات حول المجتمعات الإحصائية مثل التوزيع الطبيعي وتجانس التباين واختيار العينة من المجتمع عشوائياً.

والاختبارات اللامعلمية بصورة عامة أكثر قوة من الاختبارات المعلمية، إذ أن الإحصاءات المعلمية تميل أكثر من الإحصاءات اللامعلمية لرفض الفرضية الصفرية، وان الاختبارات اللامعلمية أسهل في طريقة إجراؤها.

الاختبارات اللامعلمية هي أساليب متحررة من التوزيع Distribution Free Methods وأساليب متحررة من الافتراضات Assumption Free Methods

الإحصاءات اللامعلمية قائمة على افتراضات ضعيفة، بينما الإحصاءات المعلمية قائمة على افتراضات قوية وأساسية، فإذا تحققت الافتراضات فإن الإحصاءات المعلمية أكثر فاعلية من

الإحصاءات اللامعلمية، وإذا لم تتحقق الافتراضات فإن الإحصاءات اللامعلمية أكثر فاعلية من الإحصاءات المعلمية.

الاسلوب اللامعلمي باختصار هو اسلوب احصائي يمتلك خصائص معينة مرغوبة تحت افتراضات ضعيفة نسبياً تخص المجتمعات المعنية التي منها تم الحصول على البيانات.

* متى نستخدم الاحصاءات اللامعلمية

1. الفرضية المراد اختبارها لا تتعلق بمعالم المجتمع.
2. البيانات المتوفرة مقاسة وفق تدريج أضعف من متطلب الاساليب المعلمية.
3. الافتراضات الخاصة بالاستخدام الملائم للاساليب المعلمية غير متحقق (انتهاك واضح لافتراضات الاساليب المعلمية).
4. الحصول على نتائج بسرعة وسهولة وتكلفة أقل.

2-12 الطرق اللامعلمية (عينة واحدة)
Nonparametric Methods (Single Sample)

1. اختبار الإشارة Sign test
2. اختبار ويلككسون للرتب ذات الإشارة Wilcoxon signed ranks test
3. اختبار السلاسل Runs test for randomness
4. اختبار كوكس ستيوارت للاتجاه Cox-Stuart test for trend
5. اختبار ذات الحدين Binomial test
6. اختبار كولمقروف سميرنوف Kolmogrov-Smirnov Test

Analyze | Nonparametric Tests ▸
Chi-Square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

12-2-1 اختبار الاشارة (عينة واحدة) Sign test (One sample)

*** الافتراضات Assumptions**

1. العينة مختاره عشوائياً من مجتمع وسيطه غير معلوم.
2. المتغير المقصود مقاس على مقياس رتبي على الأقل.
3. توزيع الاختبار الاحصائي هو توزيع ذات الحدين $B(n, ½)$

*** الفرضية المطلوب اختبارها Hypothesis**

الفرضية الصفرية: الوسيط = مقدار ثابت $H_0: \theta = \theta_0$

الفرضية البديلة: الوسيط ≠ مقدار ثابت $H_1: \theta \neq \theta_0$

*** الاختبار الاحصائي Test Statistic**

الاختبار الاحصائي هو العدد الأقل للإشارات الموجبة أو السالبة

T= r(smallest no of + or – signs)

*** اتخاذ القرار Decision Rule**

نرفض الفرضية الصفرية اذا كانت القيمة المحسوبة $P(T\leq r)$ أقل من القيمة الحرجة $\alpha/2$

نقبل الفرضية الصفرية اذا كانت القيمة المحسوبة $P(T\leq r)$ أكبر من القيمة الحرجة $\alpha/2$

مثال12-1: اختيرت عينة عشوائية تتكون من 20 كتاباً وكانت اعداد الصفحات لهذه الكتب كما يلي:

153, 166, 181, 192, 244, 248, 258, 264, 266, 305, 305, 312, 330, 340, 356, 361, 395, 427, 433, 467

استخدم هذه البيانات في اختبار الفرضية الصفرية، عند مستوى دلالة $(\alpha=0.05)$.

$H_0: \theta = 250$

$H_1: \theta \neq 250$

3-12 الطرق اللامعلمية (عينتين مستقلتين)
Nonparametric Methods (Two Independent Samples)

1. اختبار توكي السريع Tukey's Quick Test (TQT)
2. اختبار الوسيط Median Test
3. اختبار مان-وتني Mann-Whitney Test
4. اختــبار تــساوي التبايــنات – اختبار انصاري-برادلي Test for equality of variance (Ansari-Bradley Test)
5. اختــبار تــساوي التباينات – اختبار موسس Test for equality of variance (Moses Test)
6. اختــبار تــساوي التبايــنات – اختبار مربعات الرتب Test for equality of variance (The Sequare Rank Test)
7. اختبار التوزيعات المتماثلة – اختبار سميرنوف Test fir identical distribution (Smirnov Test)

*** اختبارات تساوي الوسيطات لعينتين مستقلتين:**

1. اختبار توكي السريع Tukey's Quick Test (TQT) $H_0: M_1 = M_2$
2. اختبار الوســـــيط Median Test $H_0: M_1 = M_2$
3. اختبار مــان-وتني Mann-Whitney Test $H_0: M_1 = M_2$

*** اختبارات تساوي التباينات لعينتين مستقلتين**

4. اختبار تساوي التباينات – اختبار انصاري-برادلي $H_0: \sigma^2_1=\sigma^2_2$
5. اختبار تساوي التباينات – اختبار موسس $H_0: \sigma^2_1=\sigma^2_2$
6. اختبار تساوي التباينات – اختبار مربعات الرتب $H_0: \sigma^2_1=\sigma^2_2$

*** اختبارات التماثل لعينتين مستقلتين**

التماثل: التساوي في مقاييس النزعة المركزية، التشتت، التماثل في الشكل، واخلال أي منهم يجعل التوزيعات غير متماثلة.

7. اختبار التوزيعات المتماثلة – اختبار سميرنوف $F_1(x) = F_2(x)$

اختبارات تساوي الوسيطات لعينتين مستقلتين– اختبار مان–وتني Mann-Whitney

يستخدم لاغراض مقارنة وسيطين لمجتمعين بالاعتماد على مشاهدات مستقلة تختار عشوائياً من المجتمعين.

يستخدم للمقارنة بين عينتين مستقلتين عندما تكون البيانات عددية بطبيعتها، وهو البديل اللامعلمي لاختبار T المعلمي للبيانات المستقلة.

والفرضية الصفرية له: لا يوجد اختلاف جوهري في علامات الأفراد في المجتمعات التي سحبت منها العينتان.

*** افتراضات الاختبار**

1. تتكون البيانات من عينة من المشاهدات العشوائية $x_1,x_2,...,x_{n1}$ من المجتمع الأول الذي وسيطه M1 غير معروف، ومن عينة اخرى من المشاهدات $y_1,y_2,...,y_{n2}$ من المجتمع الثاني الذي وسيطه M2 غير معروف.
2. المتغيرات متصلة x, y.
3. مقاس على سلم رتبي على الأقل.

*** الفرضية Hypothesis**

وسيط المجتمع الأول = وسيط المجتمع الثاني $H0: M1 = M2$

وسيط المجتمع الأول ≠ وسيط المجتمع الثاني $H0: M1 \neq M2$

*** لاختبار الفرضية التي تتعلق بتساوي وسيطين نقوم بالاتي:**

1. نشكل العينة الكلية من العينتين.
2. نرتب المشاهدات في العينة الكلية من الأقل الى الأكبر ونعطيها الرتب التي تبدأ من 1 إلى n1+n2 لأكبر قيمة.
3. نجمع الرتب الخاصة بمشاهدات العينة المأخوذة من المجتمع الأول ويرمز للمجموع بالرمو S.
4. نحسب الاختبار الاحصائي T كما يلي:

$$T = S - \frac{n_1(n_1+1)}{2}$$

* القرار:

رفض H_0 اذا كان قيمة الاختبار الاحصائي T اكبر من القيمة الحرجة $W_1 - (\alpha/2)$

اذا كان قيمة الاختبار الاحصائي T أقل من القيمة الحرجة $W(\alpha/2)$

حيث نجد القيمة الحرجة $W(\alpha/2)$ من الجدول A7

وتكون $W_1 - (\alpha/2) = n1n2 - W(\alpha/2)$

* **ملاحظـــة**: اذا كانـــت أي منهما أو كليهما أكبر من 20 نستخدم التقريب للتوزيع الطبيعي.

$$Z = \frac{T - (n1n2/2)}{\sqrt{n1n2\,(n1+n2+1)/12}}.$$

مثال2-12: في دراسة حول تقييم الصدق التمييزي لاختبار ما، تم تقسيم مجموعة من الأفراد الى مجموعتين متمايزتين في السمة التي يقيسها الأختبار، حيث يمتلك أفراد المجموعة الأولى والـــبالغ عددهم 17 فرداً السمة بدرجة عالية، بينما يمتلك أفراد المجموعة الثانية والبالغ عددهم 10 افـــرادً السمة بدرجة منخفضة، إذا طبق الاختبار على افراد المجموعتين وكانت العلامات لأفراد المجموعتين كما في الجدول التالي:

هل يمتلك الاختبار قدرة تمييزية عند مستوى دلالة ($\alpha = 0.05$).

* **الحل**:

	MARK-G1	MARK-G2	Sort-G1	Sort-G2	Rank-G1	Rank-G2
	11.90	6.60	11.90	6.60	27.00	15.00
	11.70	5.80	11.70	5.80	26.00	13.00
	9.50	5.40	9.50	5.40	25.00	12.00
	9.40	5.10	9.40	5.10	24.00	11.00
	8.70	5.00	8.70	5.00	23.00	9.50
	8.20	4.30	8.20	4.30	22.00	8.00
	7.70	3.90	7.70	3.90	21.00	5.00
	7.40	3.30	7.40	3.30	19.50	4.00
	7.40	2.40	7.40	2.40	19.50	3.00
	7.10	1.70	7.10	1.70	18.00	1.00
	6.90		6.90		17.00	
	6.80		6.80		16.00	
				6.60		
	6.30		6.30		14.00	
				5.80		
				5.40		
				5.10		
				5.00		
	5.00		5.00		9.5	
				4.30		
	4.20		4.20		7	
	4.10		4.10		6	
				2.4		
	2.20		2.20		2	
				1.7		
count	17	10				
sum					296.50	81.50

مجموع الرتب للعينة الأولى S1

S1=27+26+25+24+23+22+21+19.5+18+17+16+14+9.5+7+6+2=296.5

$$T = S - \frac{n_1(n_1+1)}{2} = 296.5 - (18*17)/2 = 296.5-153 = 143.5$$

46 = W (α/2) نجد أن A.7 من جدول n1=10 , n2=17 , P=0.025

$W_1 - (\alpha/2) = n1n2 - W(\alpha/2) = 18*17 - 46 = 260$

*** القرار:**

رفض H_0 اذا كان قيمة الاختبار الاحصائيT اكبر من القيمة الحرجة $W_1 - (\alpha/2)$

اذا كان قيمة الاختبار الاحصائيT أقل من القيمة الحرجة $W(\alpha/2)$

اذا نرفض الفرضية الصفرية H0، وهذا يعني أن هناك قدوة تمييزية للاختبار.

مجموع الرتب للعينة الثانية S2

S2=12+13+15+11+9.5+8+5+4+3+1=81.5

$$T = S - \frac{n_2(n_2+1)}{2} = 81.5 - (10*11)/2 = 26.5$$

46 = W (α/2) نجد أن A.7 من جدول n1=10 , n2=17 , P=0.025

$W_1 - (\alpha/2) = n1n2 - W(\alpha/2) = 10*17 - 46 = 124$

القـــرار: بمـــا أن قيمة الاحصائي T=26.5 وهي خارج حدود الفترة (124 - 46) اذا نرفض الفرضية الصفرية H0

التقريب:

$$Z = \frac{T - (n1n2/2)}{\sqrt{n1n2\,(n1+n2+1)/12}} = \frac{26.5 - 85}{\sqrt{10*17\,(10+17+1)/12}} = \frac{-58.5}{19.916} = -2.937$$

القـــرار: بمـــا أن القـــيمة المحسوبة Z = -2.937 أقل من القيمة الحرجة نرفض الفرضية الصفرية

4-12 الطرق اللامعلمية (عينتين مرتبطتين)
Nonparametric Methods (Two Related Samples)

1. اختبار الاشارة لعينتين مرتبطتين Sign Test
2. اختبار ويلككسون لاشارة رتب الفروق المطلقة Wilcoxon signed ranks test
3. اختبار مكنمار لعينتين مرتبطتين Mc Nemar's test
4. اختبار مربع كاي χ^2 للاستقلالية Chi-square test of independence
5. اختبار مربع كاي χ^2 للتجانس Chi-square test of homogeniety

*** اختبارات الوسيطات لعينتين مترابطتين:**

1. اختبار الاشـــارة لعينتين مرتبطتين $H_0: M_1 = M_2$
2. اختبار ويلككسون لاشارة رتب الفروق المطلقة $H_0: M_1 = M_2$

*** اختبارات النسب لعينتين مترابطتين:**

3. اختبار مكنمار لعينتين مرتبطتين $H_0: P_1 = P_2$

*** اختبارات الاستقلال لعينتين مترابطتين:**

4. اختبار مربع كاي χ^2 للاستقلالية

*** اختبارات التجانس لعينتين مترابطتين:**

5. اختبار مربع كاي χ^2 للتجانس

*** اختبارات الوسيطات لعينتين مترابطتين:**

*** اختبار ويلككسون لإشارة رتب الفروق المطلقة Wilcoxon signed ranks test**

تستخدم في التصاميم التجريبية ذات الاختبارين القبلي والبعدي، وهو البديل اللامعلمي لاختـــبار T للبـــيانات المترابطة، ومن أهم ميزات هذا الاختبار انه يختبر اتجاه الفرق بين أزواج المـــشاهدات وحجـــم هذا الفرق النسبي أيضاً، ويجب أن تكون المشاهدات رقمية، ولا يمكن استخدامه إذا كانت تصنيفية اسمية.

إذا كانـــت الفـــروق تفضل إحدى المجموعتين تكون هذه المجموعة هي الأفضل بدلالة إحصائية.

*** الافتراضات**

1. البيانات تتلخص بمجموعة من الازواج العشوائية.
2. يمثل الفرق بين العنصر الأول والثاني في الازواج المرتبة متغيراً متصلاً.
3. توزيع الفروق متماثل حول الوسيط للفرق.
4. الفروق مستقلة من خلال الاختيار العشوائي للأفراد.
5. الفروق مقاسة على الأقل في المستوى الفئوي.

*** الفرضية**

$H_0: M_D = 0 , M_1 = M_2$
$H_1: M_D \neq 0 , M_1 \neq M_2$

*** لأختبار الفرضية نعمل الآتي:**

1. احصل على الفروق $Xi - Yi$
2. اعطي رتباً للفروق المطلقة R
3. اعطي الرتب اشارات الفروق التي انتجتها
4. احسب مجموع الرتب ذات الاشارة الموجبة +T ومجموع الرتب السالبة -T

*** الاختبار الاحصائي** $T = K = \text{smaller} (T+ \text{ or } T-)$

*** القرار**

ارفض الفرضية الصفرية اذا كانت قيمة الاختبار الاحصائي $T \leq$ القيمة الحرجة C التي يمكن الحصول عليها من جدول A.3 ، أو ارفض الفرضية الصفرية اذا كان $P(T=K|n) \leq (\alpha/2)$ وفيما عدا ذلك نقبل الفرضية الصفرية.

* ملاحظة

عندما يكون من غير الممكن استخدام جدول A.3 بسبب أن n > 30 نستخدم القيمة التقريبية من التوزيع الطبيعي المعياري للاختبار والتي تحسب من المعادلة التالية:

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2*n+1)}{24}}}.$$

n: عدد الازواج.

ويتم مقارنة القيمة الناتجة مع القيمة الحرجة 1.96

* مثال**12-3**:

درس العالمــــان لاتاني وكابل أثر اجتماع (التقاء) مجموعة من الفئران على معدل دقات القلــــب لها وقد رصدوا هذه المعدلات لعشرة فئران بيضاء عندما كان كل منها لوحده وعندما كان مع فار آخر. ويبين الجدول التالي تلك المعدلات:

R	R_R	Di	|Di|	Rank|Di|	رتب مع الاشارة
402.00	437.00	-35.00	35	7	-7
409.00	470.00	-61.00	61	9	-9
415.00	408.00	7.00	7	2	2
418.00	448.00	-30.00	30	5	-5
426.00	454.00	-28.00	28	4	-4
450.00	476.00	-26.00	26	3	-3
456.00	535.00	-79.00	79	10	-10
462.00	461.00	1.00	1	1	1
462.00	494.00	-32.00	32	6	-6
463.00	523.00	-60.00	60	8	-8

$H_0: M_D = 0 , M_1 = M_2$

الفرضية الصفرية :لا يوجد أثر لوضع الفئران مع بعضها على دقات القلب

$H_1: M_D \neq 0 , M_1 \neq M_2$

الفرضية البديلة : يوجد أثر لوضع الفئران مع بعضها على دقات القلب

* الاختبار الاحصائي

$$T = K = \text{smaller} (T+ \text{ or } T-) = \Sigma r+ = 1+2 = 3$$
$$P(T=3 \mid n=10) = 0.0049$$

* القرار

ارفض الفرضية الصفرية اذا كانت قيمة الاختبار الاحصائي T ≥ القيمة الحرجة C التي يمكن الحصول عليها من جدول A.3

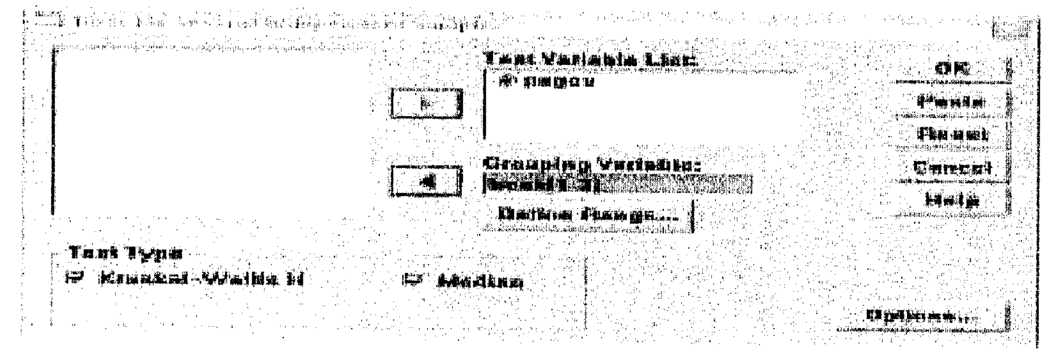
أو ارفض الفرضية الصفرية اذا كان $P(T=K|n) \leq (\alpha/2)$

بما أن $P(T=3 \mid n=10) = 0.0049$ أقل من $(\alpha/2) = 0.025$ لذلك نرفض الفرضية الصفرية، ونستنتج أن هناك أثر لوضع الفئران معاً على دقات قلبها.

5-12 الطرق اللامعلمية (ثلاثة عينات مستقلة أو أكثر)
Nonparametric Methods(3 or more Independent Samples)

1. اختبار كروسكال والس Kruskall-Wallis test
2. اختبار العلامات الطبيعية البديلة A normal scores alternative test
3. اختبار جونكير في حالة الفرض البديل المرتب Jonckheere test for ordered alternative
4. اختبار تجانس التباين Test for equality of variance

1. اختبار كروسكال والس $H_0: M_1 = M_2 = M_3$
2. اختبار العلامات الطبيعية البديلة $H_0: M_1 = M_2 = M_3$
3. اختبار جونكير في حالة الفرض البديل المرتب $H_0: M_1 = M_2 = M_3$
4. اختبار تجانس التباين $H_0: \sigma^2_1 = \sigma^2_2 = \sigma^2_3$



*** اختبار كروسكال والس Kruskall-Wallis Test**

هـــذا الاختـــبار يناظر تحليل التباين الأحادي One-Way ANOVA في الاحصاءات المعلمية، وهو تحليل تباين احادي للرتب، وهو امتداد مباشر لأختبار ويلككسون واختبار مان-وتني.

وهو مصمم لاختبار الفرضية الصفرية التي تنص على عدة عينات مستقلة قد سحبت من نفـــس المجتمع، لذا فهو مصمم لاختبار دلالة الفرق بين ثلاث مجموعات مستقلة أو أكثر عندما تكون البيانات على المتغير التابع رتبية أو يمكن ترتيبها.

Sample1	Sample2	....	Sample I	....	Sample k
X_{11}	X_{12}				
X_{21}	X_{22}				
.					
.					
X_{n11}	X_{n22}				

n_1: حجم العينة الأولى.

n_2: حجم العينة الثانية.

n_k: حجم العينة k

S_1: مجموع رتب العينة الأولى.

S_2: مجموع رتب العينة الثانية. S_k: مجموع رتب العينة K.

*** الفرضية الصفرية:**

$H_0: M_1 = M_2 = M_3$
$H_1: M_1 \neq M_2 \neq M_3$

*** الطريقة:**

1. لدينا n من القيم نرتبها من أقل قيمة لأكبر قيمة (العينة الكلية).
2. نعطيها رتب من اقل قيمة رتبة 1 واعلى قيمة الرتبة n ، وتعطى القيم المتساوية نفس الرتبة.
3. تعطى الرتب لمواقع المشاهدات في عيناتها.
4. نجمع الرتب الخاصة بكل عينة Si وهي مجموع الرتب للعينة i

5. نربع مجاميع الرتب لكل عينة $S^2_t = \Sigma \frac{S^2_i}{n_i} = \frac{S^2_1}{n_1} + \frac{S^2_2}{n_2} + ... + \frac{S^2_k}{n_k}$

6. في حالة عدم وجود ties

$$S^2_r = \Sigma r^2_{ij} = 1^2 + 2^2 + 3^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{6}$$

*** الاختبار الاحصائي**

بوجود عدد قيم متعادلة (رتب متساوية) وهي وجود ties

$$T = \frac{(n-1)\,(S^2_t - c)}{S^2_r - c}$$

$$\text{Where } c = 0.25 * n\,(n+1)^2$$

عدم وجود عدد قيم متعادلة (رتب متساوية) وهي وجود ties

$$T = \frac{12}{n\,(n+1)}\; S^2t - 3\,(n+1)$$

* **القرار:** اذا كان عدد العينات 3 والحجوم فيها 5 أو أقل نستخدم جدول A.12

المحسوبة < الحرجة نفشل في ان نرفض الفرضية الصفرية.

المحسوبة > الحرجة نرفض الفرضية الصفرية.

* **ملاحظة:** في الحالة التي لا نستطيع فيها استخدام الجدول A.12 حيث (عدد العينات أكثــر من 3 أو حجم العينات في أي منها أكثر من 5) نتوقع أ تتوزع قيم الاختبار T حسب توزيع $\chi 2$ بدرجات حرية تساوي k-1 حيث (k عدد العينات). لذلك يمكن استخدام جدول $\chi 2$ في هذه الحالة لمعرفة القيم الحرجة.

مثال 12-4:

تم حــصر عدد الصفحات في ثلاثة مجموعات عشوائية من الكتب (احصاء، رياضيات، حاســوب) واعتماداً على اعداد الصفحات المسجلة نرغب باختيار الفرضية الصفرية التي تنص علــى أن العينات الثلاث من نفس التوزيع، مقابل الفرضية البديلة التي تنص على أن هناك على الأقل عينة واحدة جاءت من توزيع له وسيط مختلف عن الوسيطات في التوزيعات الأخرى عند مستوى الدلالة الاحصائية $\alpha = 0.05$ نفذ هذا الاختبار اذا كانت عدد الصفحات كما يلي:

	Stat	Math	Comp	Stat_R	Math_R	Comp_R
	126	93	29	10	6	1
	142	98	39	11	7	2
	156	216	60	12	15	3
	228	249	78	17	20	4
	245	301	82	18	25	5
	246	319	112	19	26	8
	370	731	125	28	34	9
	419	910	170	29	36	13
	433		192	30		14
	454		224	31		16
	478		263	32		21
	503		275	33		22
			276			23
			286			24
			369			27
			756			35
مجموع الرتب				270	169	227
حجم العينة				12	8	16

K= 3 S3=227 S2= 169 S1= 270

N= 36 n3= 16 n2= 8 n1= 12

H0: الفرضية الصفرية: مجموعات الكتب الثلاثة جاءت من نفس التوزيع

H1: الفرضية البديلة: على الأقل مجموعة واحدة جاءت من توزيع له وسط مختلف

$$S^2_t = \Sigma \frac{S^2_i}{n_i} = \frac{S^2_1}{n_1} + \frac{S^2_2}{n_2} + \frac{S^2_3}{n_3} = \frac{(227)^2}{16} + \frac{(169)^2}{8} + \frac{(270)^2}{12} = 12865.7$$

$$T = \frac{12}{n(n+1)} S^2t - 3(n+1) = \frac{12}{36*37} *12865.7 - 3(37) = 4.91$$

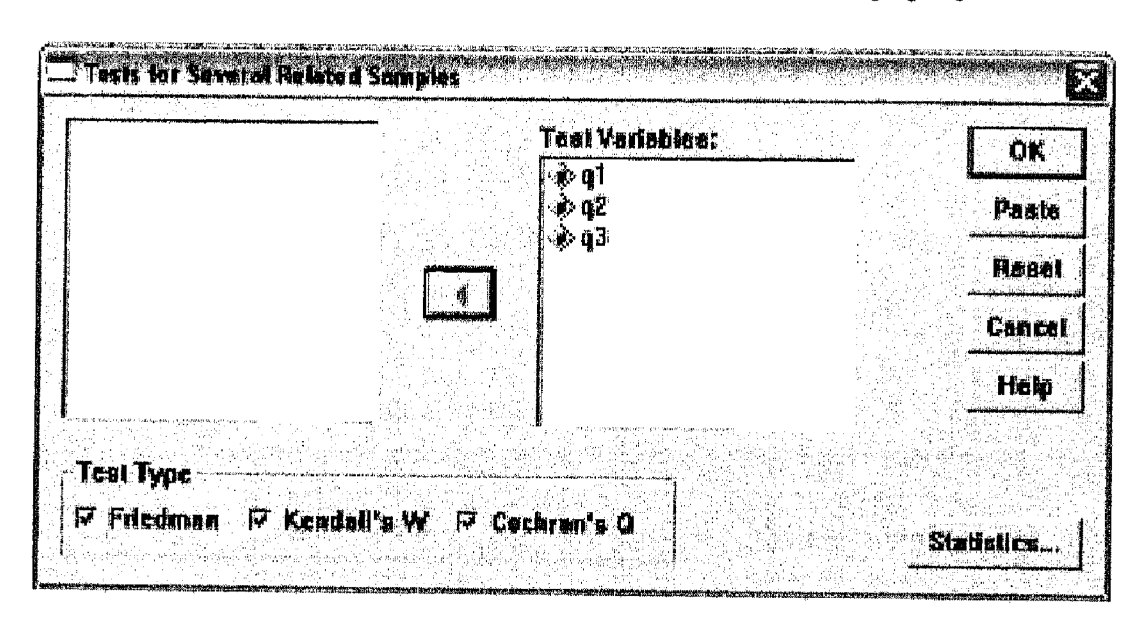
*** القرار**

بمـــا أن الحجوم لبعض العينات تزيد عن 5 فإن توزيع الاختبار الاحصائي T يقترب من توزيـــع $\chi2$ بدرجات حرية $(3-1=2)$ لذلك نفشل في رفض الفرضية الصفرية لأن المحسوبة (4.91) < الحـــرجة $\chi2_{2,0.95} = 5.991$ ونـــستنتج أن مجموعات الكتب الثلاثة من نفس

الـــتوزيع مـــن حيث عدد صفحات الكتاب، أي لها نفس الوسيطات ولها نفس التشتت ونفس الشكل.

6-12 الطرق اللامعلمية (ثلاثة عينات مرتبطة أو أكثر)
Nonparametric Methods (3 or more Related Samples)

1. اختبار فريدمان Friedman test — $H_0: M_1= M_2 = M_3$
2. اختبار كوكران Cochran's test — $H_0: M_1= M_2 = M_3$



* اختبار كوكران **Cochran's test**

يـــستخدم هـــذا الاختبار في حالة البيانات الثنائية (متغير ثنائي) مثل النجاح والفشل، الحـــدوث وعدم الحدوث، الربح والخسارة، الذكور والاناث، الحياة والموت، والبيانات المرتبطة يمكن صياغتها كمصفوفة.

الأفراد	A1	A2	... Ak	
S1				B1=ΣXij
S2				B2
.				.
Sn				
				N=ΣXij

$Tj = \Sigma Xij$

$Bi = \Sigma Xij$

لاختبار الفرضية الصفرية

جميع المعالجات متكافئة في فعاليتها :H0

هناك على الأقل معالجة تختلف في فعاليتها عن بقية المعالجات :H1

نستخدم اختبار كوكران والذي يعطى بالعلاقة التالية:

$$Q = \frac{k(k-1)\,\Sigma Tj^2 - (k-1)N^2}{KN - \Sigma Bi^2} \sim \chi^2_{k-1}$$

القرار:

المحسوبة < الحرجة نفشل في ان نرفض الفرضية الصفرية.

المحسوبة > الحرجة نرفض الفرضية الصفرية.

مثال12-5:

اذا كانت علامات 5 أشخاص في ثلاثة اسئلة موضوعية كما يلي:

الأفراد	Q1	Q2	Q3	Total
S1	1	1	0	2
S2	1	0	1	2
S3	0	0	1	1
S4	0	1	1	2
S5	1	0	1	2
	3	2	4	9

هل تدل هذه النتائج على اختلاف صعوبة الفقرات عند مستوى الدلالة الاحصائية $\alpha = 0.05$

الفرضية الصفرية: جميع الفقرات لها نفس درجة الصعوبة :H0

الفرضية البديلة: هناك على الأقل فقرة تختلف في صعوبتها عن بقية الفقرات :H1

$$Q = \frac{k(k-1)\,\Sigma Tj^2 - (k-1)N^2}{KN - \Sigma Bi^2} \sim$$

$$Q = \frac{3(3-1)[32+22+42] - 2*9^2}{3*9 - [2^2+2^2+1^2+2^2+2^2]} = \frac{174 - 162}{27 - 17} = 1.2$$

$\chi^2_{2,0.05} = 5.991$ القيمة الحرجة

القرار:

بما ان القيمة المحسوبة (1.2) < أقل من القيمة الحرجة ($\chi^2_{2,0.05} = 5.991$) نفشل في رفض الفرضية الصفرية، يعني الفقرات الثلاثة متساوية في درجة صعوبتها.

7-12 استخدام برمجية SPSS في حل المسائل.

مثال 12-2: نستخدم اختبار مان-وتني – عينتين مستقلتين.

* ادخال البيانات

group	mark
1	7.40
1	7.40
1	7.10
1	6.90
1	6.80
1	6.30
1	5.00
1	4.20
1	4.10
1	2.20
2	6.60
2	5.80
2	5.40
2	5.10
2	5.00
2	4.30
2	3.90
2	3.30
2	2.40
2	1.70

Analyze – Nonparametric Test – 2 Independent Samples…

ضع المتغير mark في نافذة: Test Variable List، والمتغير group في نافذة Grouping Variable:

من Test Type أنقر مربع الاختبار Mann-Whitney U، ثم انقر زر Ok تظهر الشاشة أدناه:

Mann-Whitney Test

Ranks

	GROUP	N	Mean Rank	Sum of Ranks
MARK	1.00	17	17.44	296.50
	2.00	10	8.15	81.50
	Total	27		

Test Statistics[b]

	MARK
Mann-Whitney U	26.500
Wilcoxon W	81.500
Z	-2.938
Asymp. Sig. (2-tailed)	.003
Exact Sig. [2*(1-tailed Sig.)]	.002[a]

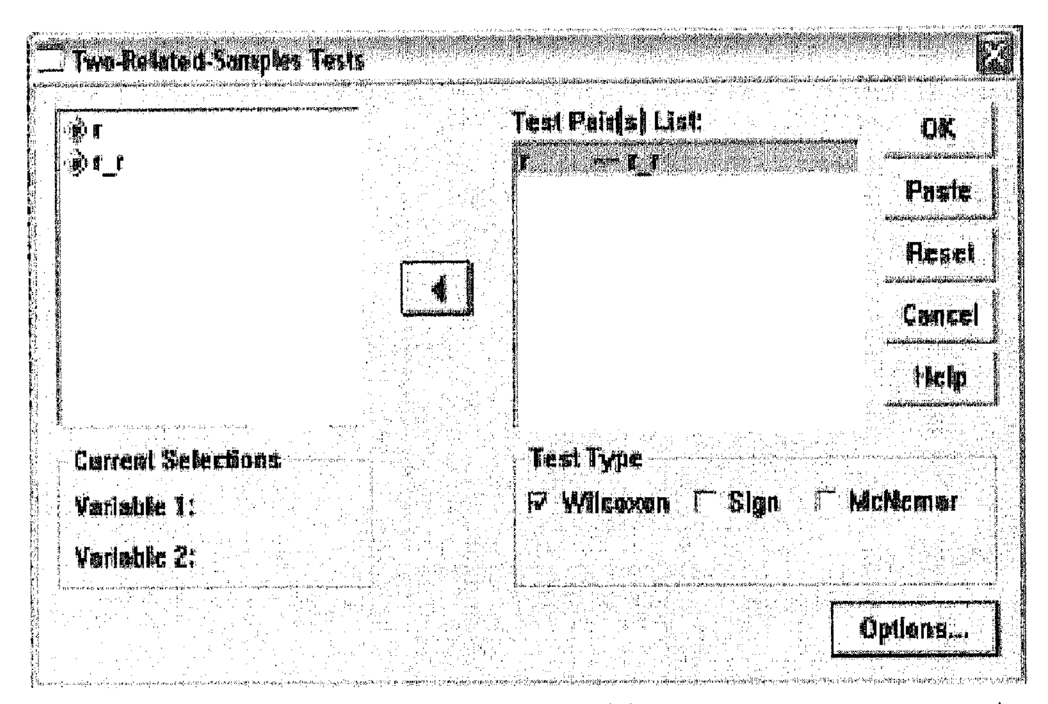
a. Not corrected for ties.

مثال 12-3: نستخدم اختبار ويلككسون – عينتين مترابطتين.

* ادخال البيانات

402	437
409	470
415	408
418	448
426	454
450	476
456	535
462	461
462	494
463	523

Analyze – Nonparametric Test – 2 Related Samples...



ضع المتغيران r , r_r في نافذة Test Pair(s) List:

من Test Type أنقر مربع الاختبار Wilcoxon، ثم انقر زر Ok تظهر الشاشة أدناه:

NPar Tests

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
R_R - R	Negative Ranks	2[a]	1.50	3.00
	Positive Ranks	8[b]	6.50	52.00
	Ties	0[c]		
	Total	10		

a. R_R < R

b. R_R > R

c. R = R_R

Test Statistics[b]

	R_R - R
Z	-2.497[a]
Asymp. Sig. (2-tailed)	.013

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

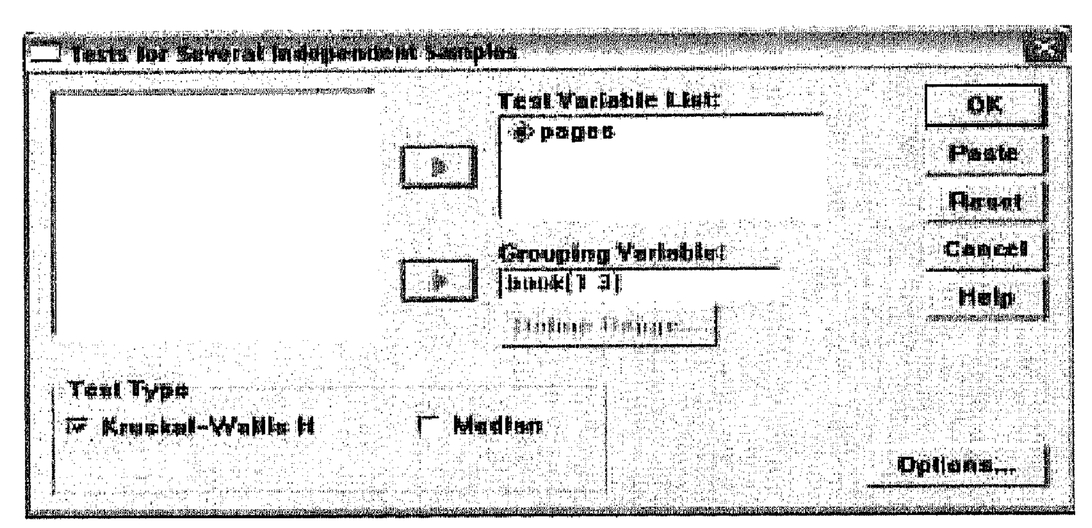
مثال **4-12**: نستخدم اختبار كروسكال والس – ثلاث عينات مستقلة أو أكثر

* ادخال البيانات

	book	pages
1	1.00	126.00
2	1.00	142.00
3	1.00	156.00
4	1.00	228.00
5	1.00	245.00
6	1.00	246.00
7	1.00	370.00
8	1.00	419.00
9	1.00	433.00
10	1.00	454.00
11	1.00	478.00
12	1.00	503.00
13	2.00	93.00
14	2.00	98.00
15	2.00	216.00
16	2.00	249.00
17	2.00	301.00
18	2.00	319.00
19	2.00	731.00
20	2.00	910.00

21	3.00	29.00
22	3.00	39.00
23	3.00	60.00
24	3.00	78.00
25	3.00	82.00
26	3.00	112.00
27	3.00	125.00
28	3.00	170.00
29	3.00	192.00
30	3.00	224.00
31	3.00	263.00
32	3.00	275.00
33	3.00	276.00
34	3.00	286.00
35	3.00	369.00
36	3.00	756.00

Analyze - Nonparametric Test - K Independent Samples...



ضــع المــتغير pages في نافــذة :Test Variable List، والمــتغير book في نافــذة Grouping Variable:

مــن Test Type أنقــر مربع الاختبار Kruskal-Wallis H، ثم انقر زر Ok تظهر الشاشة أدناه:

→ NPar Tests

Kruskal-Wallis Test

Ranks

	BOOK	N	Mean Rank
PAGES	stat	12	22.50
	math	8	21.13
	comp	16	14.19
	Total	36	

Test Statistics a,b

	PAGES
Chi-Square	4.907
df	2
Asymp. Sig.	.086

a. Kruskal Wallis Test

b. Grouping Variable: BOOK

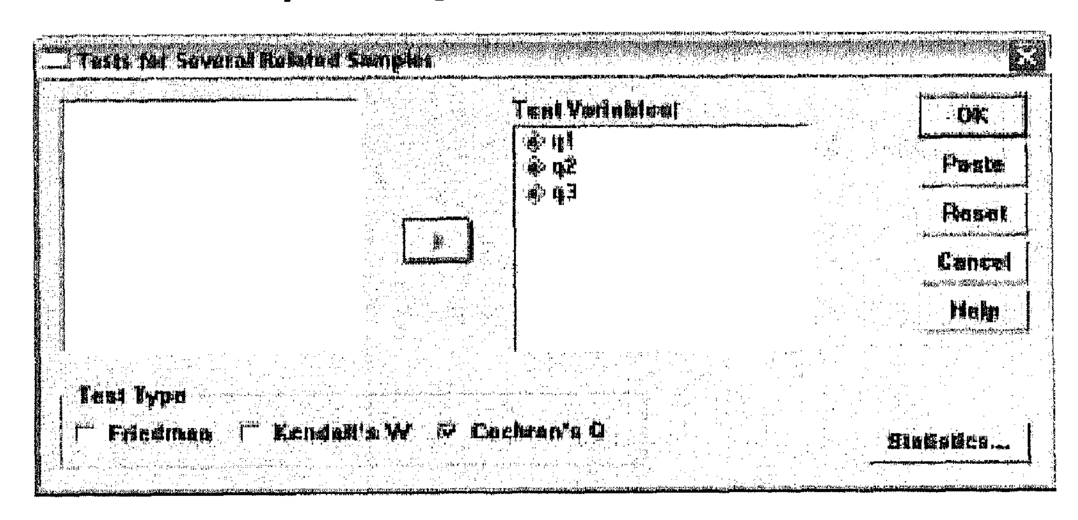
القرار: بما أن القيمة المحسوبة $\chi^2 = 4.907$ أقل من القيمة الحرجة $\chi^2_{2,0.95} = 5.991$ نفشل في رفض الفرضية الصفرية، ونستنتج أن مجموعات الكتب الثلاث من نفس التوزيع من حيث عدد صفحات الكتاب، أي أن لها نفس الوسيطات ولها نفس التشتت ونفس الشكل.

مثال 12-5: نستخدم **الطرق اللامعلمية(ثلاثة عينات مرتبطة أو أكثر)*اختبار كوكران**

* ادخال البيانات

q1	q2	q3
1	1	0
1	0	1
0	0	1
0	1	1
1	0	1

Analyze – Nonparametric Test – K Related Samples…



ضع المتغيران q1, q2, q3 في نافذة :Test Variables

منTest Type أنقر مربع الاختبارCochran's Q، ثم انقر زر Ok تظهر الشاشة أدناه:

➔ NPar Tests

Cochran Test

Frequencies

	Value	
	0	1
Q1	2	3
Q2	3	2
Q3	1	4

Test Statistics

N	5
Cochran's Q	1.200[a]
df	2
Asymp. Sig.	.549

a. 1 is treated as a success.

8-12 تمــــاريـــــن Exercise.

س1: التالـــية علامــــات 15 طالب في مساق مهارات الحاسوب قبل وبعد تلقيهم لدوره في الحاسوب، هل للدورة التدريبية أثر في تحسن هؤلاء الطلاب؟

* استخدم اختبار ويلككسون للأزواج المترابطة.

رقم الطالب	العلامة قبل الدورة	العلامة بعد الدورة	الفرق=بعدي-قبلي	الرتب للفرق	الرتب ذات الاشارة الأقل تكراراً
1	65	70			
2	40	45			
3	60	55			
4	65	82			
5	80	60			
6	25	50			
7	40	63			
8	30	50			
9	85	83			
10	75	72			
11	45	55			
12	70	70			
13	74	88			
14	35	65			
15	65	80			

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
POST - PRE	Negative Ranks	4[a]	4.38	17.50
	Positive Ranks	10[b]	8.75	87.50
	Ties	1[c]		
	Total	15		

a. POST < PRE

b. POST > PRE

c. PRE = POST

Test Statistics[b]

	POST - PRE
Z	-2.200[a]
Asymp. Sig. (2-tailed)	.028

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

س2: اراد باحث مقارنة مجموعة تجريبية مع أخرى ضابطة في ادائهما على مقياس للاتجاهات، وكانت درجاتهم كما هو مبين أدناه:

* استخدم اختبار مان-وتني لعينتين مستقلتين.

المجموعة التجريبية Expermental	Rank الرتبة	المجموعة الضابطة Control	الرتبة Rank
52		52	
68		39	
42		47	
49		38	
36		27	
31		18	
29		20	
28		15	
50			

Mann-Whitney Test

Ranks

	GROUP	N	Mean Rank	Sum of Ranks
MARK	expermental	9	10.72	96.50
	Control	8	7.06	56.50
	Total	17		

Test Statistics[b]

	MARK
Mann-Whitney U	20.500
Wilcoxon W	56.500
Z	-1.492
Asymp. Sig. (2-tailed)	.136
Exact Sig. [2*(1-tailed Sig.)]	.139[a]

a. Not corrected for ties.

b. Grouping Variable: GROUP

س3: لنفـــرض أن شـــركة قامت باستخدام ثلاثة برامج لتدريب الموظفين، وبعد الانتهاء قام باجـــراء قياسات رتبية على ادؤهم في العمل، والجدول أدناه يبين علاماتهم على المقياس المستخدم، والمطلوب معرفة إن كان هناك اختلاف فيما بين المجموعات يعزى للاسلوب التدريبي.

* استخدم اختبار كرسكال والس.

Group_1	Group_2	Group_3
75	70	60
70	60	50
80	50	45
65	40	50
		49

➔ NPar Tests

Kruskal-Wallis Test

Ranks

	GROUP	N	Mean Rank
MARK	1.00	4	11.13
	2.00	4	6.00
	3.00	5	4.50
	Total	13	

Test Statistics[a,b]

	MARK
Chi-Square	6.926
df	2
Asymp. Sig.	.031

a. Kruskal Wallis Test

b. Grouping Variable: GROUP

س4: ضع رمز الإجابة الصحيحة (نعم، لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- تتطلب الاختبارات اللامعلمية الايفاء بافتراضات تجانس التباين.

2- تتطلب الاختبارات اللامعلمية الايفاء بافتراضات التوزيع الطبيعي.

3- الاختبارات المعلمية أكثر قوة من الاختبارات اللامعلمية.

4- تميل الاختبارات اللامعلمية لرفض الفرضية الصفرية.

5- يعتبر اختبار ولكوكسون بديلاً لاختبار (T) للبيانات المستقلة.

6- يعتبر اختبار مان-وتني بديلاً لاختبار (T) للبيانات المترابطة.

7- يعتبر اختبار كروسكال والس بديلاً لتحليل التباين الاحادي.

الفصل الثالث عشر

تطبيقات البرنامج الاحصائي SPSS

1-13 التعرف على بيئة النظام الإحصائي SPSS
2-13 تشغيل نظام SPSS.
3-13 شاشات نظام SPSS.
4-13 ملفات نظام SPSS.
5-13 القوائم الرئيسية لبرنامج SPSS.
6-13 استخدام برمجية SPSS في حل المسائل.

الفصل الثالث عشر

تطبيقات البرنامج الاحصائي SPSS

13-1 التعرف على بيئة النظام الإحصائي SPSS

* أن كلمة SPSS تتكون من Statistical Package for Social Sciences وهي تعني الحزمة الإحصائية للعلوم الاجتماعية، وتستخدم لأجراء عمليات إحصائية كثيرة وبشكل سهل. وتوجد أغلب التحليلات في قائمة التحليل Analyze المبينة أدناه:

Analyze
- Reports
- Descriptive Statistics
- Compare Means
- General Linear Model
- Correlate
- Regression
- Classify
- Data Reduction
- Scale
- Nonparametric Tests
- Multiple Response

يلزمنا أن نقوم بإجراء بعض التحليلات الإحصائية للبيانات مثل: الإحصاء الوصفي Descriptive Statistics، مثل إيجاد مقاييس النزعة المركزية: الوسط الحسابي، الوسيط، المنوال، ... وإيجاد مقاييس التشتت: المدى، التباين، الانحراف المعياري، المدى المتوسط، ...

مقارنة الأوساط Compare Means.

إيجاد معاملات الارتباط Correlate: الارتباط البسيط، الارتباط المتعدد، الارتباط الجزئي، الارتباط شبه الجزئي، ....

إيجاد معادلة الانحدار Regression

الإحصاءات اللامعلمية Nonparametric Tests

إن القيام بالحسابات اليدوية لهذه الإحصاءات ليس سهلاً خاصة إذا كان حجم البيانات كبيراً، ولكن باستخدام برنامج النظام الإحصائي SPSS تصبح جميع هذه التحليلات سهلة وممتعة إذا تم التعامل مع البرنامج بشيء من الحرفية.

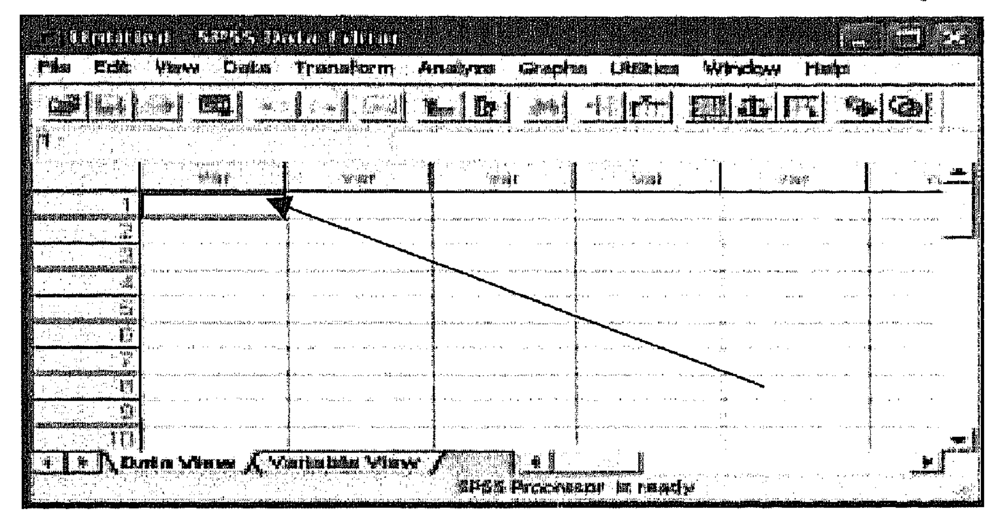
2-13 تشغيل نظام SPSS.

* تشغيل نظام **SPSS**

Start – Programs – SPSS for Windows – SPSS10.0for Windows – Type in data – Ok



تظهر لديك الشاشة المبينة أدناه:

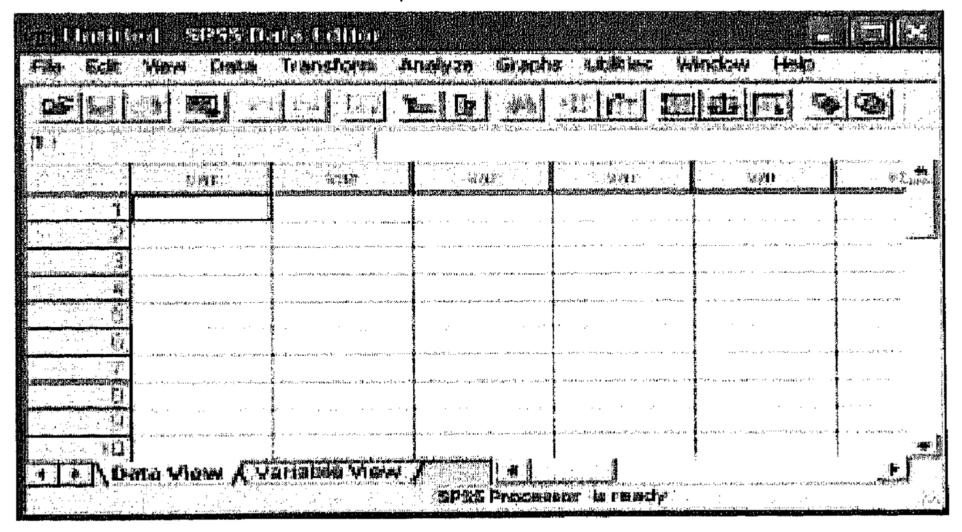


3-13 شاشات نظام SPSS:

يحتوي نظام SPSS على ثلاث شاشات رئيسية هي:

1. شاشة محرر البيانات **Data Editor Window**: وتحتوي على البيانات الإحصائية المراد تحليلها.

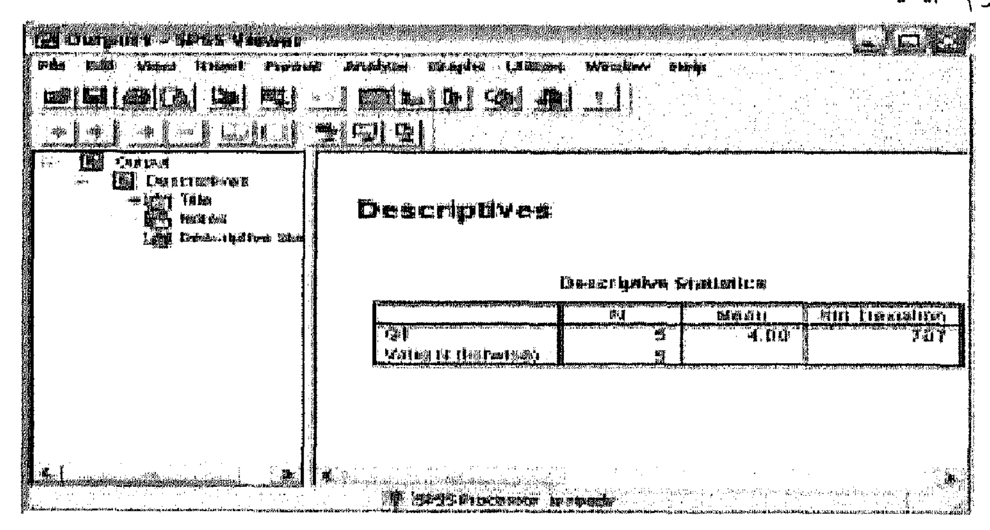
وتتكون من شاشتين هما Data View وتستخدم لإدخال البيانات



وشاشة Variable View وهي شاشة تعريف المتغيرات كما هو مبين أدناه:

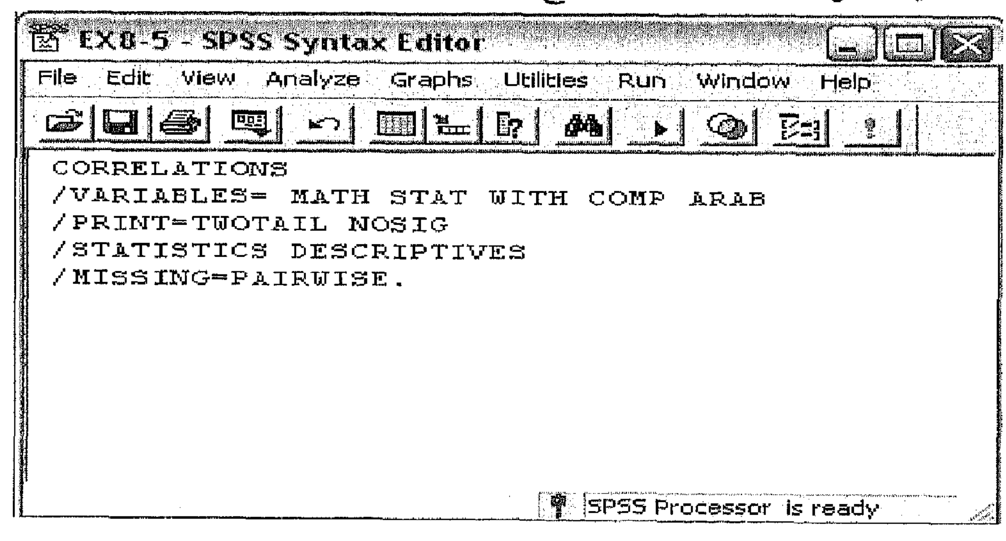
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1										
2										
3										

2. شاشة المخرجات **Output Navigator**: تظهر نتائج التحليلات الإحصائية والرسوم البيانية.

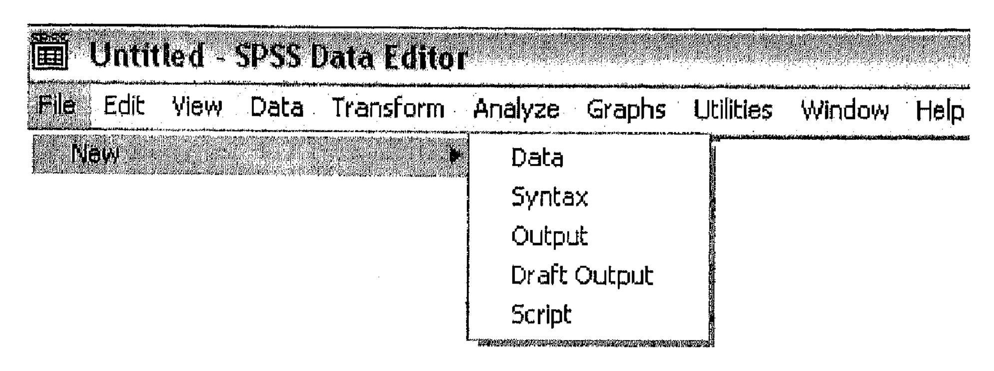


3. شاشة التعليمات **Syntax Window**: تستخدم لكتابة التعليمات للعمليات المختلفة.

من قائمة File اختر New ثم Syntax تظهر لك شاشة التعليمات أدناه، ثم اكتب التعليمات في داخل الشاشة كما هو موضح أدناه:



4-13 ملفات نظام SPSS:



هناك عدة أنواع من الملفات منها ما يلي:

1. ملـــف البيانات **Data**: وهي الملفات التي تحتوي على البيانات الخام المراد تحليلها، ويكـــون نـــوع الملف (.SAV)، فإذا كان اسم ملف البيانات JAMAL فإن نوعه (.SAV)، فيكون الاسم الكامل للملف هو (JAMAL.SAV)

2. ملـــف المخـــرجات **Output**: وهـــو الملف الذي يحتوي على نتائج الإجراءات الإحـــصائية التي تظهر في شاشة المخرجات، ويكون نوع الملف (.SPO)، فإذا كان اسم ملف المخرجات JAMAL فإن نوعه (.SPO)، فيكون الاسم الكامل للملف هو (JAMAL.SPO)

3. ملف التعليمات **Syntax**: وهو الملف الذي يحتوي على التعليمات المراد إجراؤها، ويكـــون نـــوع الملف (.SPS)، فإذا كان اسم ملف التعليمات JAMAL فإن نوعه (.SPS)، فيكون الاسم الكامل للملف هو (JAMAL.SPS)

4. ملـــف **Draft Output**: وهـــو الملف الذي يحتوي على مسودة ملف المخرجات المراد إجراؤها، ويكون نوع الملف (.rtf). مثال ذلك هو (JAMAL.rtf)

5. ملف **Script**: وهو الملف الذي يحتوي على الاجراءات الجاهزة وتستخدم لغة Sax Bacic والتي تكتب بها الاجراءات الجاهزة مثل Clean Viewer ويكون نوع الملف (.SbS)، فـــإذا كان اسم ملف الاجراءات Remove Labels فإن نوعه (.sbs)، فيكون الاسم الكامل للملف هو (Remove Labels.sbs).

5-13 القوائم الرئيسية لبرنامج SPSS:

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

قائمة ملف File

File

New
Open
Open Database
Read Text Data

Save Ctrl+S
Save As...

Display Data Info...
Apply Data Dictionary...
Cache Data...

Print... Ctrl+P
Print Preview

Switch Server...
Stop Processor Ctrl+.
Recently Used Data
Recently Used Files

Exit

New: إنشاء ملفات جديدة.

Open: فتح ملفات مخزنة سابقاً.

Read Text Data: قراءة ملف بيانات.

Save: تخزين ملفات.

Display Data Information: إظهار معلومات عن الملفات.

Print: طباعة الملفات.

Print Preview: معاينة الملفات قبل الطباعة.

Exit: الخروج من البرنامج.

قائمة تحرير Edit

Edit	
Undo Set Cell Value	Ctrl+Z
Redo	Ctrl+R
Cut	Ctrl+X
Copy	Ctrl+C
Paste	Ctrl+V
Paste Variables...	
Clear	Del
Find...	Ctrl+F
Options...	

Undo Set Cell Value: التراجع عن التحرير.

Redo Set Cell Value: التراجع عن التراجع عن التحرير.

Cut: قص البيانات.

Copy: نسخ البيانات.

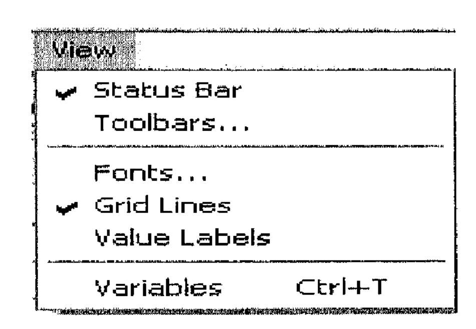
Paste: لصق البيانات.

Clear: حذف (عمود، صف) أو اكثر بما يحويه من بيانات بعد تحديد المراد حذفه.

Find: البحث عن حالات.

Options...: خيارات.

قائمة عرض View



Status Bar: إخفاء أو إظهار شريط الحالة.

Toolbars.: إظهار شريط الأدوات والأيقونات المختصرة المستخدمة بدل من القوائم.

Fonts...: تغيير نوع الخط المستخدم.

Grid Lines: إظهار أو إخفاء خطوط الشبكة.

Value Labels: إظهار أو إخفاء عناوين القيم.

Variables: إظهار شاشة تعريف المتغيرات.

قائمة البيانات Data

Data
- Define Dates....
- Insert Variable
- Insert Cases
- Go to Case...
- Sort Cases...
- Transpose...
- Restructure....
- Merge Files
- Aggregate...
- Split File.....
- Select Cases....
- Weight Cases....

Define Dates...: تعريف المتغيرات وتغيير اسماؤها.

Insert Variables: اضافة متغيرات جديدة..

Insert Cases: اضافة حالات جديدة.

Go to Case...: الذهاب الى حالة معينة.

Sort Cases...: ترتيب الملف حسب قيم متغير ما.

Transpose...: تحويل البيانات.

Merge Files: دمج الملفات وهي دمج اكثر من ملف وجعلها ملف واحد.

Aggregate...: تجميع وتلخيص الحالات.

Split File...: تقسيم الملف الى عدة أجزاء.

Select Cases...: تستخدم لاختيار مجموعة من الحالات ينطبق عليها شرط معين.

قائمة التحويلات Transform:

انشاء متغير جديد من خلال المتغيرات الموجودة.

Transform
- Compute...
- Random Number Seed...
- Count...
- Recode ▶
- Categorize Variables...
- Rank Cases...
- Automatic Recode...
- Create Time Series...
- Replace Missing Values...
- Run Pending Transforms

Compute: القيام بالعمليات الحسابية المختلفة.

Count : حساب عدد القيم المتشابهة لقائمة من المتغيرات لكل فرد من أفراد العينة.

Recode: اعادة ترميز البيانات.

Rank Cases...: إنشاء متغيرات جديدة تحتوي على رتب المتغيرات الموجودة المختلفة للقيم الرقمية.

Automatic Recode...: (الترميز الآلي)، إعادة ترميز السلاسل الحرفية إلى قيم.

Create Time Series...: إنشاء متغير جديد يحتوي متسلسلة زمنية.

Replace Missing Values...: تعويض القيم المفقودة بطرائق إحصائية.

قائمة التحليلات Analyze

Analyze
- Reports ▶
- Descriptive Statistics ▶
- Compare Means ▶
- General Linear Model ▶
- Correlate ▶
- Regression ▶
- Classify ▶
- Data Reduction ▶
- Scale ▶
- Nonparametric Tests ▶
- Multiple Response ▶

Reports: التقارير.

Descriptive Statistics : الإحصاء الوصفي.

Compare Means : مقارنة الأوساط، تحليل التباين الأحادي.

General Linear Model: تحليل التباين الثنائي Tow-Way ANOVA

Correlate : حساب معاملات الارتباط.

Regression : حساب معادلة الانحدار.

Classify : التصنيف.

Data Reduction : التحليل العاملي.

Scale : تحليل الثبات (معاملات الثبات).

Nonparametric Test : الإحصاءات اللامعلمية.

Multiple Response : تعريف المجموعات.

قائمة الرسومات البيانية Graphs

Graphs
Gallery
Interactive ▶
Bar...
Line...
Area...
Pie...
High-Low...
Pareto...
Control...
Boxplot...
Error Bar...
Scatter...
Histogram...
P-P...
Q-Q...
Sequence...
ROC Curve...
Time Series ▶

قائمة الفوائد Utilities

Utilities
Variables...
File Info
Define Sets...
Use Sets...
Run Script...
Menu Editor ...

Variables... : إعطاء معلومات عن المتغيرات.

File Info : إيجاد معلومات مفصلة عن الملف المستخدم والمتغيرات التي به.

Define Sets...: تعريف المجموعات للمتغيرات المختلفة.

Use Sets...: استخدام المجموعات للمتغيرات المختلفة.

قائمة نافذة Window

Window
Minimize All Windows
✔ 1 Untitled - SPSS Data Editor

Minimize All Windows: التحكم بحجم النوافذ.

قائمة المساعدة Help

Help
Topics
Tutorial
SPSS Home Page
Statistics Coach
About...

Topics : إعطاء مساعدة عن أي محتوى من محتويات البرنامج.

6-13 استخدام برمجية SPSS في حل المسائل.

س2) في اختبار لمادة الحاسوب الذي يتكون من الأسئلة أدناه، حلل الاختبار؟

رقم الطالب: التخصص: 1-علمي 2-أدبي 3-معلوماتية 4- مهني

الجنـــس : 1-ذكر 2-أنثى .

المادة: الحاسوب. المستوى الثالث – تراسل البيانات والشبكات. (15 سؤال).

ضع دائرة حول رمز الإجابة الصحيحة في كل مما يلي:

1- أي من التالية ليس من مكونات شبكة الحاسوب:

أ- أجهزة حاسوب مزودة ببطاقات خاصة بالاتصال. ج- منظم التيار الكهربائي.

ب- أسلاك توصيل بين الأجهزة. د- بطاقات الشبكات.

2- تغطي الشبكة المحلية LAN مساحة قطرها حوالي:

أ- 10م ب- 1000م ج-100م د- 10000م

3- الجهاز الذي يعمل على زيادة حجم وكفاءة الشبكة هو:

أ- المقسم Switch. ب- الخط Line. ج- الموزع Hub. د- الجسر Bridge.

4- شبكة مكونة من جهازي حاسوب فقط مرتبطين بخط اتصال تسمى:

أ- الشبكة المحلية LAN. ج- شبكة الخادم والمستفيد Client/Server

ب- الشبكة التناظرية Peer-to-Peer. د- الشبكة النجمية Star Network.

5- إن المصطلح Hyper Text يعني:

أ- نص مترابط. ب- عنوان الموقع. ج- مزود خدمة الانترنت. د- متصفح الانترنت.

6- إن المصطلح URL يعني:

أ- نص مترابط. ب- عنوان الموقع. ج- مزود خدمة الانترنت. د- متصفح الانترنت.

7- إن المصطلح Search Engines يعني:

أ- محركات البحث. ب- عنوان الموقع.

ج- الشبكة العنكبوتية العالمية. د- متصفح الانترنت

8- أي من التالية ليست من محركات البحث Search Engines:

أ- Yahoo ب- Google ج- Altavista د- Windows

9- يرمز للمواقع الحكومية بـــ:

أ- Edu ب- Org ج- Gov د- Net

10- تغطي الشبكة واسعة التغطية WAN مساحة تقدر بـــ:

أ- مدينة. ب- دولة. ج- قارة. د- الكرة الأرضية.

11- تقسم الشبكات المحلية LAN إلى:

أ- الشبكة التناظريةPeer-to-Peer، الشبكة التماثلية Symetric Network

ب- شبكة الخادم والمستفيد Client/Server، الشبكة التناظريةPeer-to-Peer

ج- الشبكة التناظرية Peer-to-Peer، شبكة القيمة المضافة VAN.

د- شبكة الخادم والمستفيد Client/Server، شبكة القيمة المضافة VAN.

12- تمتاز الشبكة التناظرية Peer-to-Peer بأنها:

أ- مناسبة عندما تكون شبكة الاتصال كبيرة.

ب- يوجد جهاز أساسي يؤدي الخدمات للشبكة وبقية الأجهزة محطات عمل.

ج-تكون جميع الأجهزة في هذه الشبكة متكافئة.

د- لكل جهاز حق الوصول إلى الشبكة حسب أهميته.

13- يتكون عنوان البريد الالكتروني من:

أ- اسم خاص بالمستخدم ، إشارة # ، اسم الموقع الذي يقدم خدمة البريد الالكتروني.

ب- اسم خاص بالمستخدم، إشارة $ ، اسم الموقع الذي يقدم خدمة البريد الالكتروني.

ج- اسم خاص بالمستخدم، إشارة & ، اسم الموقع الذي يقدم خدمة البريد الالكتروني.

د- اسم خاص بالمستخدم، إشارة @ ، اسم الموقع الذي يقدم خدمة البريد الالكتروني.

14- إن المصطلح WWW يعني:

أ- الشبكة العنكبوتية العالمية

ب- عنوان الموقع.

ج- محركات البحث

د- متصفح الانترنت.

15- تصنف شبكات الحاسوب حسب المنطقة الجغرافية التي تغطيها إلى:

أ- شبكات محليةLAN، شبكات واسعة التغطيةWAN.

ب- شبكة الخادم والمستفيد Client/Server، الشبكة التناظريةPeer-to-Peer

ج-شبكات محليةLAN، شبكة القيمة المضافة VAN.

د- شبكات واسعة التغطية WAN، شبكة القيمة المضافة VAN.

*** الترميز Coding**

عملية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

مثال: متغير الجنس Sex وهو (ذكر، أنثى)، حيث يعطى: الرقم 1 للذكور و2 للإناث.

- ذكر. 1
- أنثى. 2

التخصص: يعطى 1-علمي 2-أدبي 3-معلوماتية 4- مهني كما يلي:

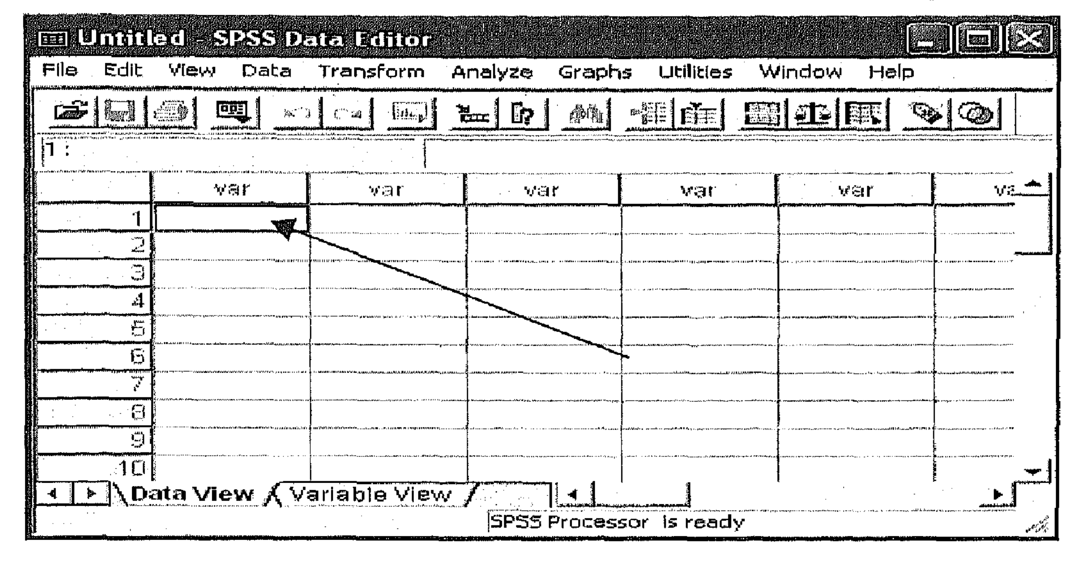
- علمي. 1
- أدبي. 2
- معلوماتية. 3
- مهني. 4

رقم الطالب، الأسئلة:

مثال1: انشيء ملف جديد اسمه example1

File – New - Data

تظهر لديك الشاشة المبينة أدناه:



لعمل جدول ترميز للمتغيرات الواردة في المثال يتم كما يلي:

انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1										
2										
3										

إدخال متغير رقم الطالب stno

إدخـــال مـــتغير الجنس sex وعند وصولنا إلى خانة Valuesوهي القيم المحتملة للمتغير نعبأها بالشكل التالي:

Value Labels

Value Labels

Value: 2

Value Label: انثى

Add | 1 = "ذكر"

Change

Remove

OK

Cancel

Help

إدخال متغير التخصص spec وعند وصولنا إلى خانة Valuesوهي القيم المحتملة للمتغير نعبأها بالشكل التالي:

Value Labels

Value Labels

Value:

Value Label:

Add | Change | Remove

1 = "علمي"
2 = "أدبي"
3 = "معلوماتية"
4 = "مهني"

OK | Cancel | Help

إدخال متغيرات الأسئلة من سؤال 1 إلى سؤال 15 وهي q1 ... q15

وبعد الانتهاء من تعريف المتغيرات تظهر الشاشة كما هو مبين أدناه:

	Name	Type	Width	Decimals	Label	Values	Missing	Column	Align	Measu
1	stno	Numeric	3	0		None	None	3	Center	Scale
2	sex	Numeric	1	0		{1, ذكر}	None	8	Right	Scale
3	spec	Numeric	1	0		{1, علمي}	None	8	Right	Scale
4	q1	Numeric	1	0		None	None	3	Center	Scale
5	q2	Numeric	1	0		None	None	3	Center	Scale
6	q3	Numeric	1	0		None	None	3	Center	Scale
7	q4	Numeric	1	0		None	None	3	Center	Scale
8	q5	Numeric	1	0		None	None	3	Center	Scale
9	q6	Numeric	1	0		None	None	3	Center	Scale
10	q7	Numeric	1	0		None	None	3	Center	Scale
11	q8	Numeric	1	0		None	None	3	Center	Scale
12	q9	Numeric	1	0		None	None	3	Center	Scale
13	q10	Numeric	1	0		None	None	3	Center	Scale
14	q11	Numeric	1	0		None	None	3	Center	Scale
15	q12	Numeric	1	0		None	None	3	Center	Scale
16	q13	Numeric	1	0		None	None	3	Center	Scale
17	q14	Numeric	1	0		None	None	3	Center	Scale
18	q15	Numeric	1	0		None	None	3	Center	Scale

إدخال البيانات Input Data

	stno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1
21	21	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
22	22	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
23	23	2	2	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1
24	24	2	2	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0
25	25	2	3	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1
26	26	2	3	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
27	27	2	3	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1
28	28	2	3	1	1	0	1	0	0	1	0	1	1	1	0	1	1	1
29	29	2	3	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
30	30	2	3	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

حفظ البيانات

File - Save As... -

Save Data As

File name: example1

Variables...

Save

Paste

Cancel

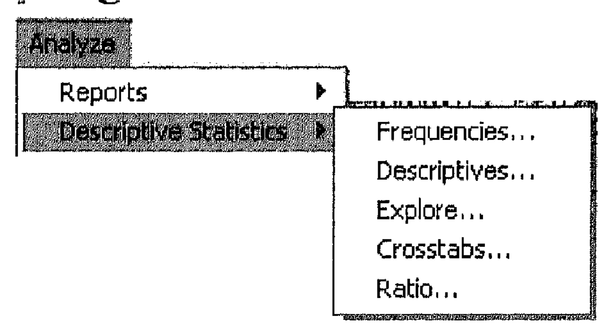
نكتب اسم الملف example1 ثم ننقر زر Save فيحفظ الملف بهذا الاسم.

التحليلات المطلوبة:

1- الإحصاء الوصفي Descriptive Statistics

Descriptive Statistics - Frequencies Analyze -

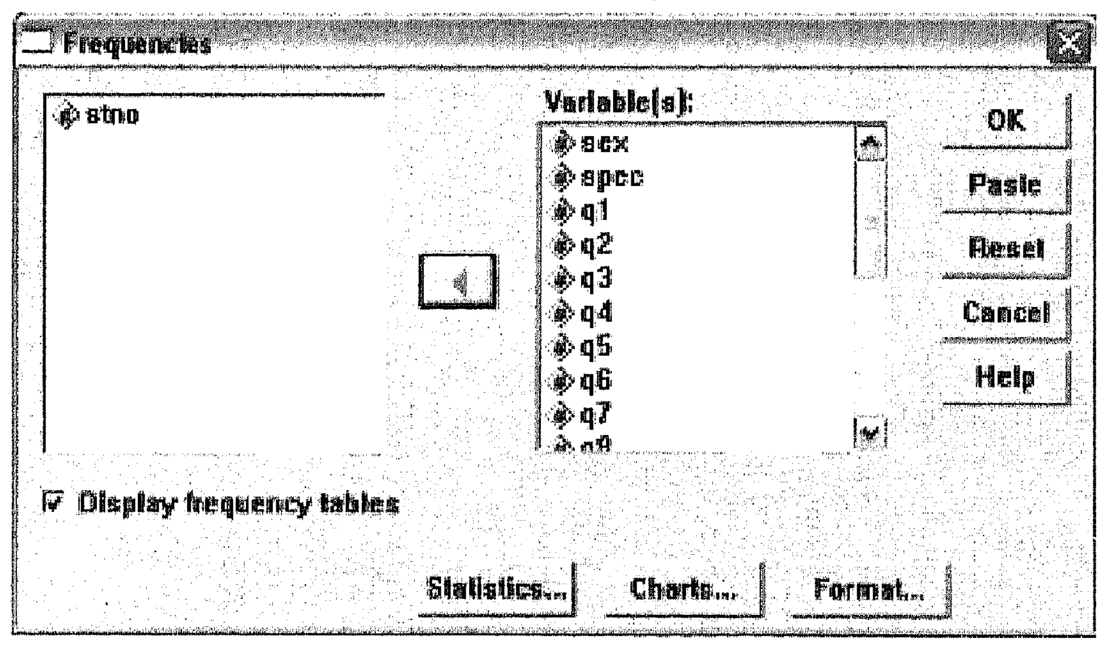
Frequencies: وصف توزيع أفراد العينة حسب أحد المتغيرات من النوع الاسمي.



بعض أسئلة الدراسة:

س1: ما نسبة الذكور والإناث في عينة الدراسة؟ (متغير الجنس sex)

س2: ما عدد أفراد العينة في كل فئة من فئات التخصص؟ (متغير التخصص spec)



بالضغط على زر ...Statistics (الاحصاءات) تظهر الشاشة المبينة أدناه:

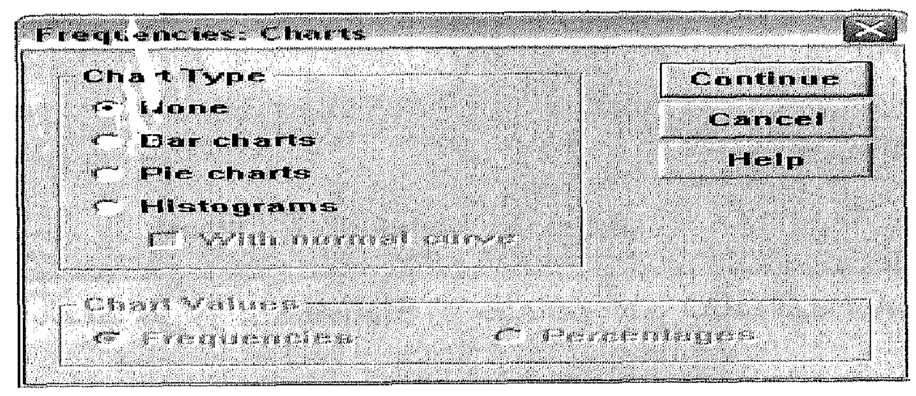
يـــستخدم لاستخراج بعض إحصاءات مقاييس النزعة المركزية Central Tendency: مثل Mean الوسط الحسابي، Median الوسيط، Mode المنوال، Sum المجموع.

يستخدم لاستخراج بعض احصاءات مقاييس التشتت: مثل Std. deviation الانحراف المعياري، Variance التباين، Range المدى.

يـــستخدم لاســـتخراج بعـــض احـــصاءات القيم المئينية: مثل Quartiles الربيعات، Percentiles المئينات.

يـــستخدم لاســـتخراج بعـــض احـــصاءات شكل التوزيع: مثل Skewness الالتواء، Kurtosis التفرطح.

بالضغط على زر ...Charts (الرسومات البيانية) تظهر الشاشة المبينة أدناه:



نوع الرسم البياني: Chart Type

1- الاعمدة Bar charts

2- الدوائر Pie charts

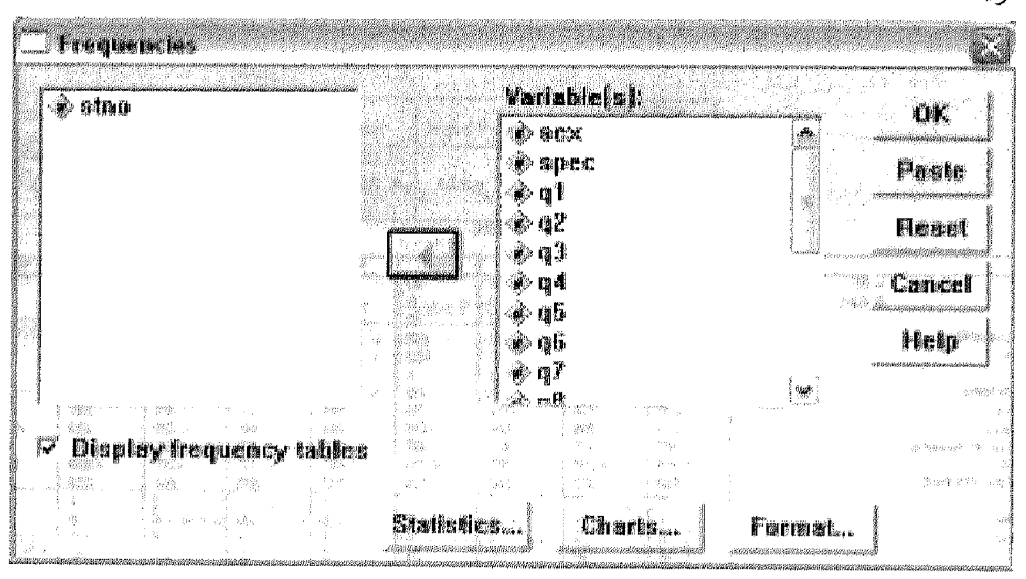
3- الرسم البياني مع المنحنى الطبيعي Histogram

* لإيجاد الوسط Mean، الوسيط Median، المنوال Mode، المجموع Sum

Analyze - Descriptive Statistics - Frequencies...



نعمل على اختيار المتغيرات وهي sex,spec,q1...q15 ثم نضغط زر Statistics ونحدد ما نريده



نحدد Mean, Median, Mode, Sum ثم نضغط زر Continue

Frequencies: Statistics

Percentile Values	Central Tendency	
☑ Quartiles	☑ Mean	Continue
☐ Cut points for [ ] equal groups	☑ Median	Cancel
☑ Percentile(s): [ ]	☑ Mode	Help
Add / Change / Remove — 50	☑ Sum	
	☐ Values are group midpoints	

Dispersion		Distribution
☑ Std. deviation	☑ Minimum	☑ Skewness
☑ Variance	☑ Maximum	☑ Kurtosis
☑ Range	☑ S.E. mean	

تظهر المخرجات المبينة أدناه:

الاحصاءات المطلوبة لكل متغير من متغيرات الدراسة:

Statistics

		sex	spec	q1	q2	q3	q4	q5	q6	q7
N	Valid	30	30	30	30	30	30	30	30	30
	Missing	0	0	0	0	0	0	0	0	0
Mean		1.57	1.27	.83	.67	.17	.53	.77	.43	.83
Std. Error of Mean		.092	.082	.069	.088	.069	.093	.078	.091	.069
Median		2.00	2.00	1.00	1.00	.00	1.00	1.00	.00	1.00
Mode		2	1	1	1	0	1	1	0	1
Std. Deviation		.504	.450	.379	.479	.379	.507	.430	.504	.379
Variance		.254	.722	.144	.230	.144	.257	.185	.254	.144
Skewness		-.315	.203	-1.884	-.745	1.884	-.141	-1.322	.283	-1.884
Std. Error of Skewness		.427	.427	.427	.427	.427	.427	.427	.427	.427
Kurtosis		-2.054	-4.233	1.657	-1.554	1.657	-2.127	-.271	-2.063	1.657
Std. Error of Kurtosis		.833	.833	.833	.833	.833	.833	.833	.833	.833
Range		8	2	1	1	1	1	1	1	1
Minimum		8	1	0	0	0	0	0	0	0
Maximum		2	3	1	1	1	1	1	1	1
Sum		47	38	25	20	5	16	23	13	25
Percentiles	25	1.00	1.00	1.00	.00	.00	.00	.75	.00	1.00
	50	2.00	2.00	1.00	1.00	.00	1.00	1.00	.00	1.00
	75	2.00	2.00	1.00	1.00	.00	1.00	1.00	1.00	1.00

		q8	q9	q10	q11	q12	q13	q14	q15
N	Valid	30	30	30	30	30	30	30	30
	Missing	0	0	0	0	0	0	0	0
Mean		.83	.97	.80	.67	.67	.87	.93	.93
Std. Error of Mean		.069	.033	.074	.088	.088	.063	.046	.046
Median		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Mode		1	1	1	1	1	1	1	1
Std. Deviation		.379	.183	.407	.479	.479	.346	.254	.254
Variance		.144	.033	.166	.230	.230	.120	.064	.064
Skewness		-1.884	-5.477	-1.580	-.745	-.745	-2.273	-3.660	-3.660
Std. Error of Skewness		.427	.427	.427	.427	.427	.427	.427	.427
Kurtosis		1.657	30.000	.527	-1.554	-1.554	3.356	12.207	12.207
Std. Error of Kurtosis		.833	.833	.833	.833	.833	.833	.833	.833
Range		1	1	1	1	1	1	1	1
Minimum		0	0	0	0	0	0	0	0
Maximum		1	1	1	1	1	1	1	1
Sum		25	29	24	20	20	26	28	28
Percentiles	25	1.00	1.00	1.00	.00	.00	1.00	1.00	1.00
	50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	75	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

توزيع أفراد العينة حسب متغيري الجنس sex، التخصص spec، الاسئلة q1...q15

Frequency Table

SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ذكر	10	33.3	33.3	33.3
	انثى	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

SPEC

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	علمي	11	36.7	36.7	36.7
	أدبي	9	30.0	30.0	66.7
	معلوماتية	10	33.3	33.3	100.0
	Total	30	100.0	100.0	

Q1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	16.7	16.7	16.7
	1	25	83.3	83.3	100.0
	Total	30	100.0	100.0	

Q2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	33.3	33.3	33.3
	1	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

Q3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	25	83.3	83.3	83.3
	1	5	16.7	16.7	100.0
	Total	30	100.0	100.0	

Q4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	14	46.7	46.7	46.7
	1	16	53.3	53.3	100.0
	Total	30	100.0	100.0	

Q5

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	7	23.3	23.3	23.3
	1	23	76.7	76.7	100.0
	Total	30	100.0	100.0	

Q6

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	18	60.0	60.0	60.0
	1	12	40.0	40.0	100.0
	Total	30	100.0	100.0	

Q7

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	16.7	16.7	16.7
	1	25	83.3	83.3	100.0
	Total	30	100.0	100.0	

Q8

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	16.7	16.7	16.7
	1	25	83.3	83.3	100.0
	Total	30	100.0	100.0	

Q9

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1	3.3	3.3	3.3
	1	29	96.7	96.7	100.0
	Total	30	100.0	100.0	

Q10

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6	20.0	20.0	20.0
	1	24	80.0	80.0	100.0
	Total	30	100.0	100.0	

Q11

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	33.3	33.3	33.3
	1	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

Q12

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	33.3	33.3	33.3
	1	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

Q13

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	4	13.3	13.3	13.3
	1	26	86.7	86.7	100.0
	Total	30	100.0	100.0	

Q14

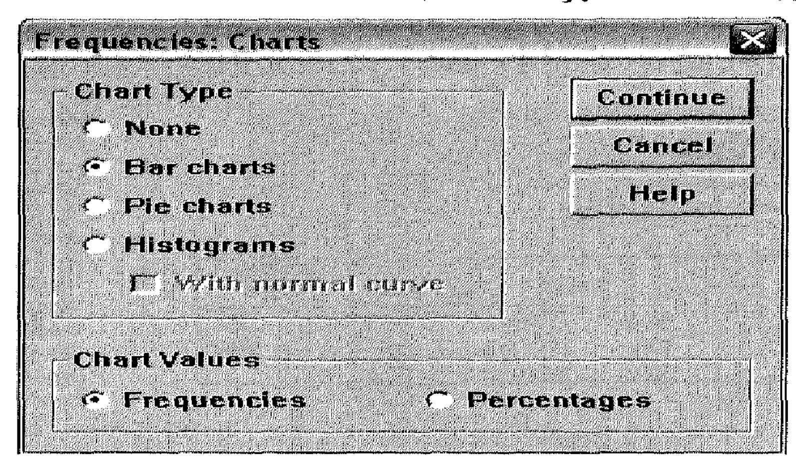
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	2	6.7	6.7	6.7
	1	28	93.3	93.3	100.0
	Total	30	100.0	100.0	

Q15

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	2	6.7	6.7	6.7
	1	28	93.3	93.3	100.0
	Total	30	100.0	100.0	

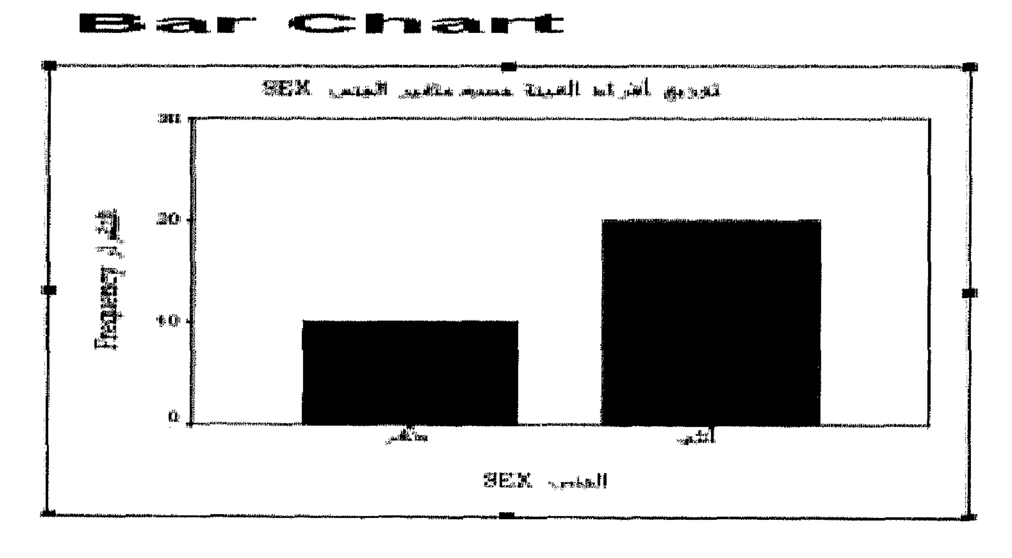
تمثيل النتائج باستخدام الرسم البياني

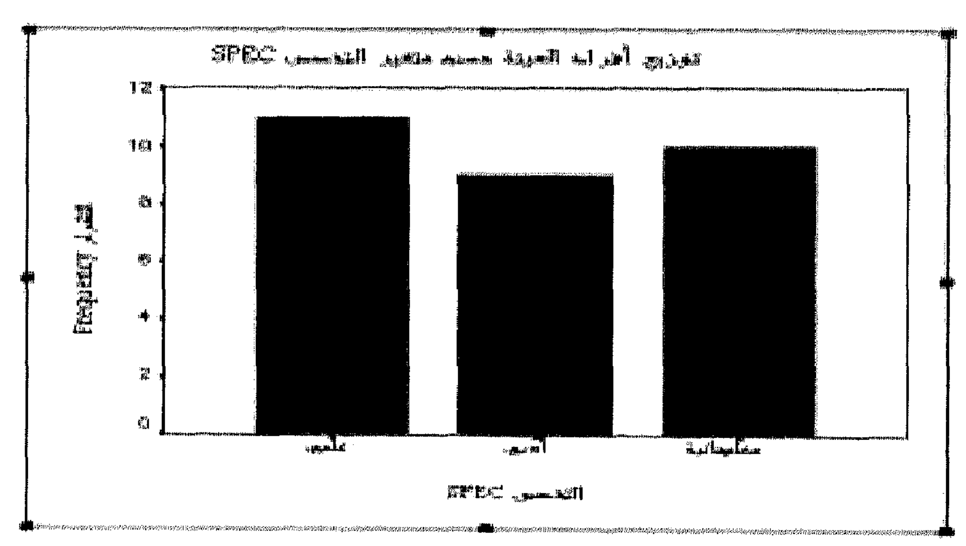
بالضغط على زر ...Charts تظهر الشاشة المبينة أدناه:



نختار نوع Bar charts ثم نضغط زر Continue ثم نضغط زر Ok

Bar Chart





النتائج:

تتكون العينة من (30) فرداً، عدد الذكور (10) طالباً، وعدد الاناث (20) طالبة.

تتكون العينة من (30)فرداً، الفرع العلمي(11)،الفرع الأدبي(9)، فرع المعلوماتية(10)

Frequency Table

SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ذكر	10	33.3	33.3	33.3
	انثى	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

SPEC

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	علمي	11	36.7	36.7	36.7
	أدبي	9	30.0	30.0	66.7
	معلوماتية	10	33.3	33.3	100.0
	Total	30	100.0	100.0	

بالنسبة للسؤال الأول، اجاب عليه (25) اجابة صحيحة، واجاب عليه(5) اجابة خاطئة.

بالنسبة للسؤال الثاني،اجاب عليه(20) اجابة صحيحة، واجاب عليه(10) اجابة خاطئة.

Q1

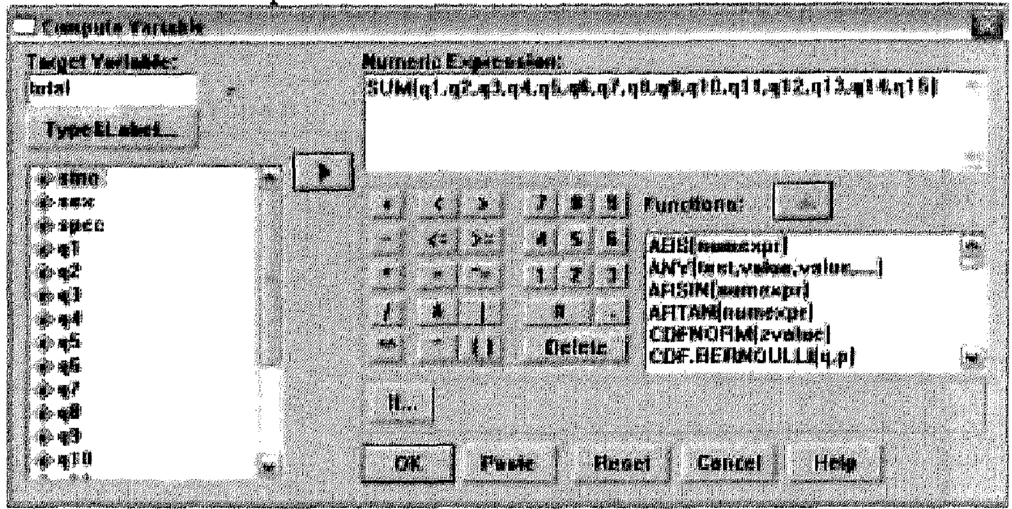
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	16.7	16.7	16.7
	1	25	83.3	83.3	100.0
	Total	30	100.0	100.0	

Q2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	33.3	33.3	33.3
	1	20	66.7	66.7	100.0
	Total	30	100.0	100.0	

* لإيجاد متغير اسمه total وهو مجموع علامات كل طالب على الاختبار

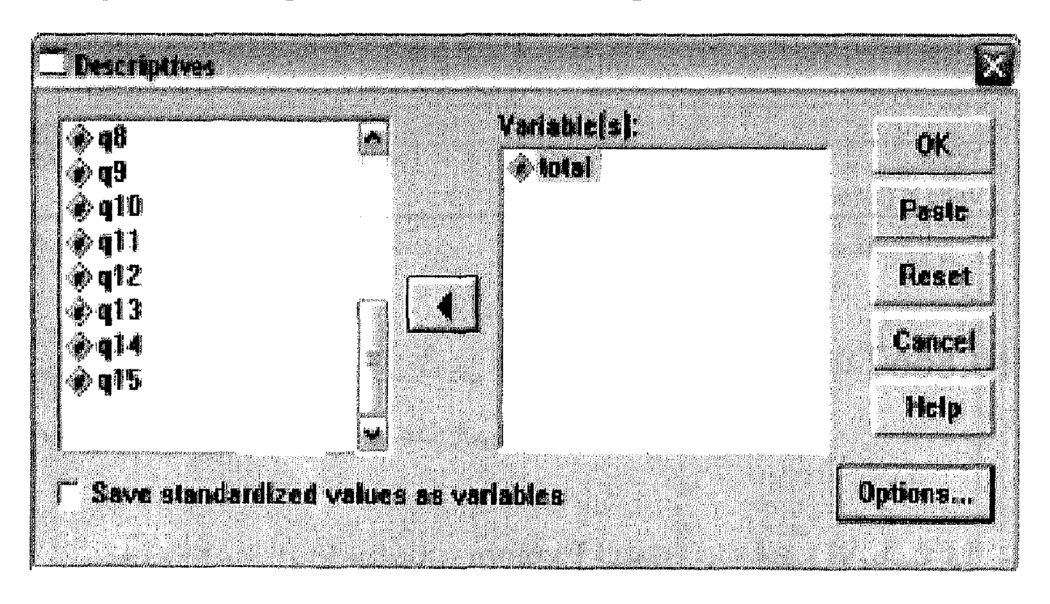
Transform – Compute...



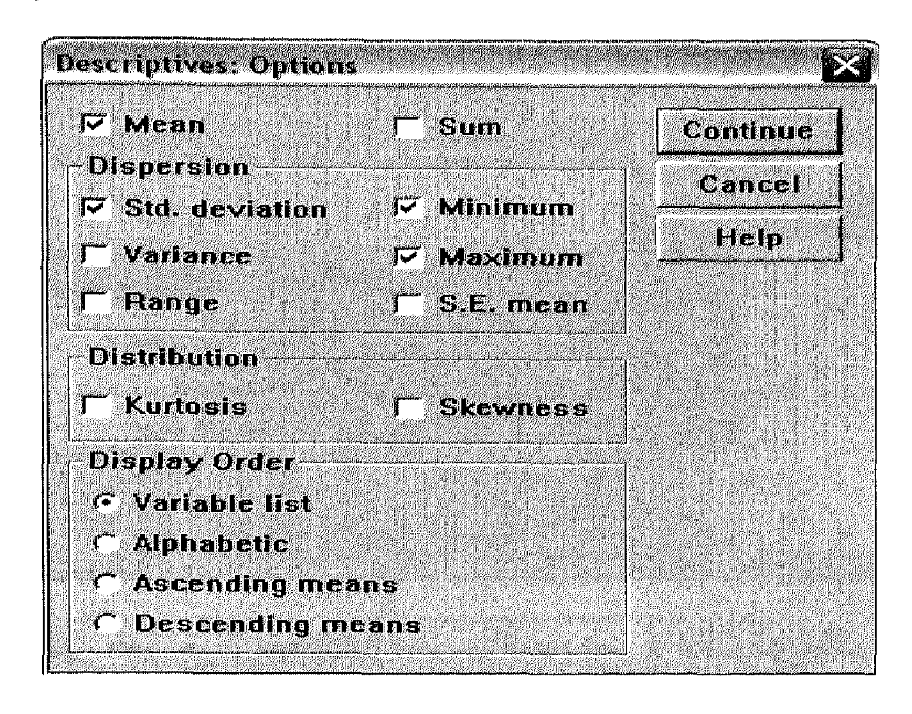
يظهر متغير جديد اسمه total وهو عبارة عن مجموع علامة كل طالب في جميع الاسئلة من 1- 15 باستخدام الاقتران Sum من قائمة Functions:

* لإيجاد الوسط Mean، الوسيط Median، المنوال Mode، المجموع Sum

Analyze – Descriptive Statistics – Descriptives...



اضغط زر ...Options تظهر لديك الشاشة المبينة أدناه:



اختر الاحصاءات المطلوبة ثم إضغط زر Continue ثم Ok تظهر النتائج أدناه:

➜ Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
TOTAL	30	3.00	14.00	10.8667	3.00268
Valid N (listwise)	30				

النتيجة:

تـــراوحت علامات الطلاب بين 3 إلى 14، وبلغ المتوسط الحسابي لها 10.8667 بانحراف معياري 3.00268 ، وكان عدد العينة 30 شخصاً.

* يـــستخدم الاجـــراء الاحـــصائي Explore للتحقق من الخطوة الأساسية قبل إجراء التحلـــيلات الاحصائية وهي فحص البيانات ومحاولة تصحيح الاخطاء إن وجدت مثل الارقام غـــير المنطقـــية أو الـــشاذة أو التحقق من أن توزيع المتغير طبيعياً، أو التحقق من شرط تجانس التباين.

* يستخدم إجراء **Explore** لعمل ما يلي:

حساب الاحصاءات الوصفية.

عمل بعض الرسومات لتوضيح شكل توزيع المتغيرات مثل Histograms, Stem-and-Leaf Plot, Box Plot

اختبار التوزيع الطبيعي عن طريق اختبار Shapiro Wilks واختبار Lilliefors

اختبار تجانس التباين Homogeneity of Variances عن طريق اختبار Levene-Test

* حساب الاحصاءات الوصفية للمتغير total

Analyze – Descriptive Statistics – Explore...

Descriptives

			Statistic	Std. Error
TOTAL	Mean		10.8667	.54821
	95% Confidence Interval for Mean	Lower Bound	9.7454	
		Upper Bound	11.9879	
	5% Trimmed Mean		11.0926	
	Median		12.0000	
	Variance		9.016	
	Std. Deviation		3.00268	
	Minimum		3.00	
	Maximum		14.00	
	Range		11.00	
	Interquartile Range		3.2500	
	Skewness		-1.050	.427
	Kurtosis		.384	.833

* استخدام اختبار **CHI-Square**

الجدول المتقاطع–مربع كاي **Crosstabulation Table: Chi-Square**

أن مربع كاي χ^2 من الإحصائيات الهامة ولها عدة استخدامات منها الكشف عن عملية الاستقلالية Independence بين متغيرين عندما تكون هناك تكرارات ويكون لكل متغير عدة مستويات محدودة.

خطوات عمل الجدول المتقاطع–مربع كاي

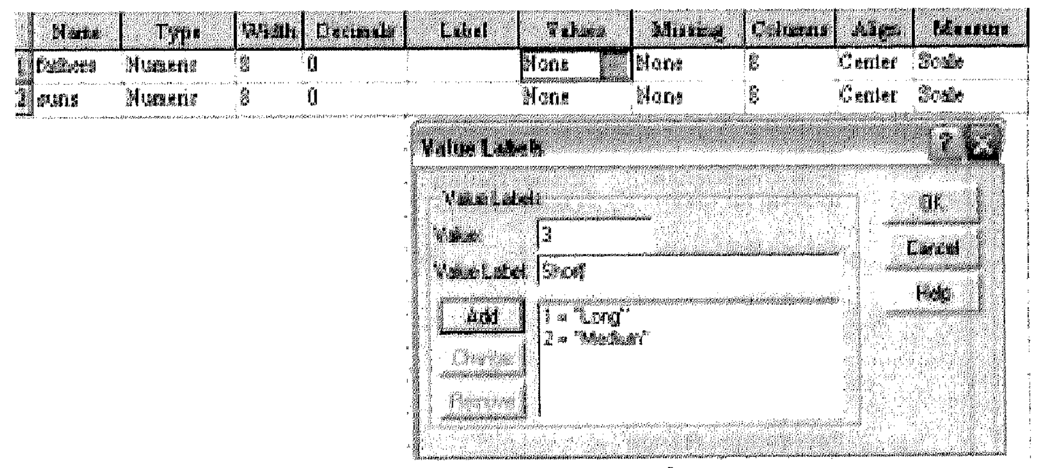
مـــثال: البيانات التالية تمثل الأطوال للآباء (Fathers) ويمثل ثلاث مستويات (طويل، متوسط، قصير) لمجموعة من الآباء والأطوال لمجموعة من الأبناء (Suns) ويمثل ثلاث مستويات (طويل، متوسط، قصير) والتي تشمل 20 أب مع أبنائهم.

المطلوب: هل يوجد ارتباط بين المتغيرين باستخدام مربع كاي χ^2.

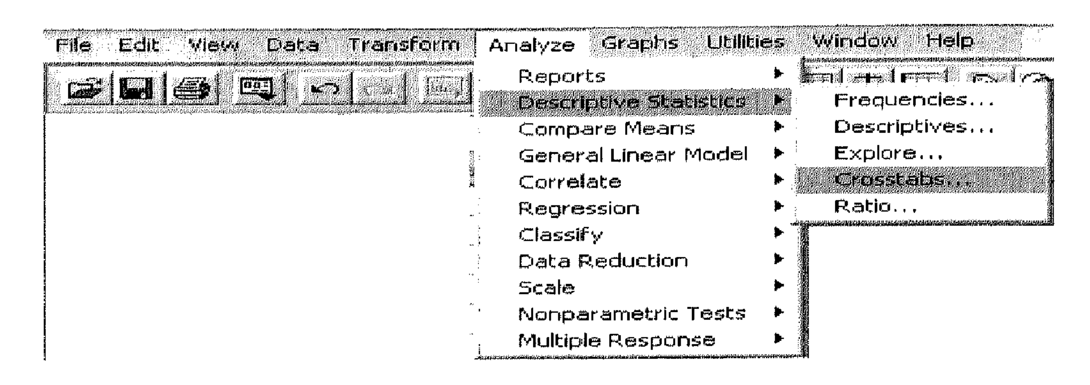
1- تعريف المتغيرات وهي:

Father ولها ثلاث مستويات: (Long , Medium , Short)

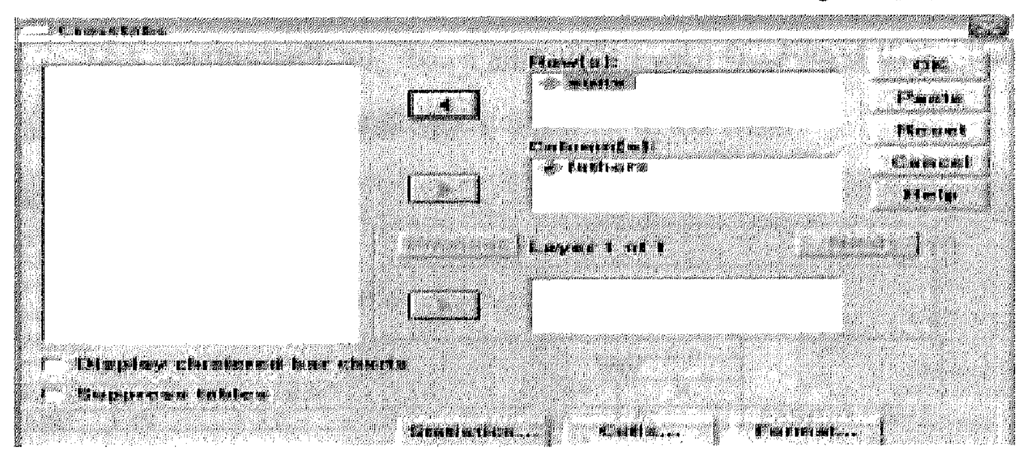
Suns ولها ثلاث مستويات: (Long , Medium , Short)



Analyze – Descriptive Statistics – Crosstabs…



- نضع المتغير التابع (Suns) في حقل الصفوف Rows ونضع المتغير المستقل (Fathers) في حقل الأعمدة Columns



نضغط على زر Statistics... لتحديد نوع مربع كاي المطلوب Cells...

ونضع اشارتي صح بجانب Chi-square والارتباط Correlations وبعدها نضغط على زر الاستمرار Continue

نضغط على زر Cells... ونحدد الخيارات المطلوبة وهي:

Crosstabs: Cell Display

Counts
- [x] Observed
- [x] Expected

Continue
Cancel
Help

Percentages
- [ ] Row
- [x] Column
- [ ] Total

Residuals
- [ ] Unstandardized
- [ ] Standardized
- [ ] Adj. standardized

من الخيار Counts نحدد التكرار المشاهد Observed والتكرار المتوقع Expected

مـــن الخيار Percentage نختار العمود Columns ويعطي النسبة المئوية للمتغير المستقل

وبعدها نضغط على زر الاستمرار **Continue** ثم نضغط زر الموافقة **OK**

تظهر شاشة المخرجات أدناه:

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SUNS * FATHERS	20	100.0%	0	.0%	20	100.0%

SUNS * FATHERS Crosstabulation

			FATHERS			Total
			Long	Medium	Short	
SUNS	Long	Count	5	2	0	7
		Expected Count	2.1	2.8	2.1	7.0
		% within FATHERS	83.3%	25.0%	.0%	35.0%
	Medium	Count	1	6	0	7
		Expected Count	2.1	2.8	2.1	7.0
		% within FATHERS	16.7%	75.0%	.0%	35.0%
	Short	Count	0	0	6	6
		Expected Count	1.8	2.4	1.8	6.0
		% within FATHERS	.0%	.0%	100.0%	30.0%
Total		Count	6	8	6	20
		Expected Count	6.0	8.0	6.0	20.0
		% within FATHERS	100.0%	100.0%	100.0%	100.0%

يمـــثل نتائج التكرارات ونسبها في الخلايا والتي تم اضافتها في خيار Cell فقد تم اضافة expected وكذلك percentage .

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26.667[a]	4	.000
Likelihood Ratio	29.439	4	.000
Linear-by-Linear Association	14.794	1	.000
N of Valid Cases	20		

a. 9 cells (100.0%) have expected count less than 5. The minimum expected count is 1.80.

السطر الأول من الجدول هو المطلوب والمهم حيث أن قيمة مربع كاي Pearson Chi-Square تساوي (26.667) وهذا يعني أن هناك علاقة قوية وموجبة بين أطوال الآباء والأبناء حيث أن مستوى المعنوية يساوي 0.000 .

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Interval by Interval	Pearson's R	.882	.067	7.957	.000[c]
Ordinal by Ordinal	Spearman Correlation	.878	.083	7.795	.000[c]
N of Valid Cases		20			

a. Not assuming the null hypothesis.

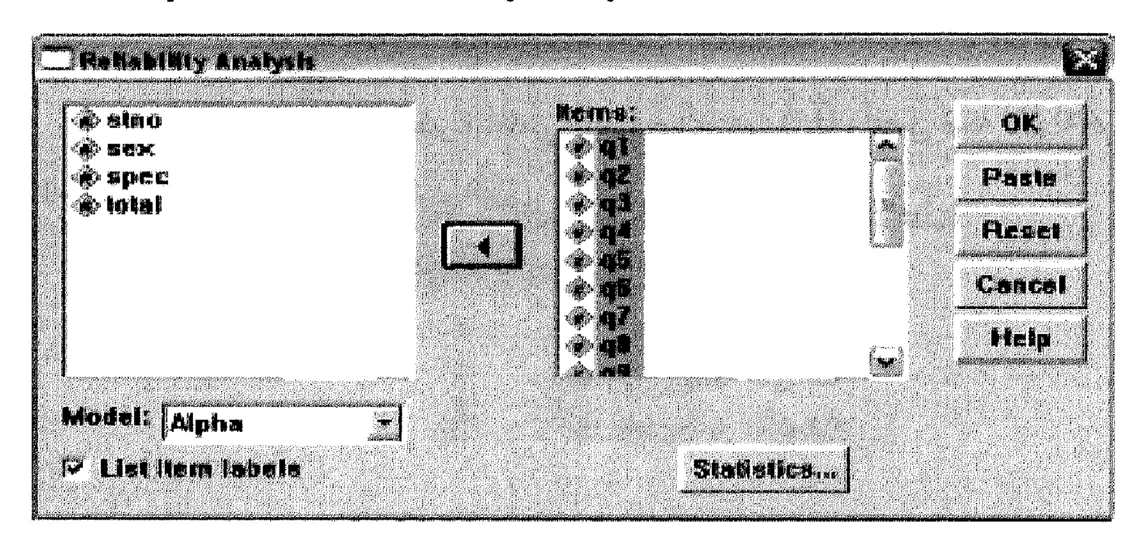
b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

* الثبات **Reliability** : اختبار كرونباخ الفا Cronbach Alpha

يستخدم هذا الاختبار لتحديد ثبات الاختبار

Analyze – Scale – Reliability Analysis...



نختار المتغيرات المطلوبة وهي الاسئلة من q1-q15 ونضعها في قائمة Items، ثم نضغط زر Ok، فتظهر شاشة المخرجات أدناه:

Reliability

```
****** Method 1 (space saver) will be used for this analysis ******
 R E L I A B I L I T Y   A N A L Y S I S -  S C A L E   (A L P H A)
  1.     Q1
  2.     Q2
  3.     Q3
  4.     Q4
  5.     Q5
  6.     Q6
  7.     Q7
  8.     Q8
  9.     Q9
 10.     Q10
 11.     Q11
 12.     Q12
 13.     Q13
 14.     Q14
 15.     Q15
Reliability Coefficients
N of Cases =      30.0                    N of Items = 15

Alpha =      .7860
```

وتكـــون اصـــغر قـــيمة مقبولة لمعامل كرونباخ الفا هي 0.6 وأفضل قيمة مقبولة هي (0.7-0.8) وكلما زادت القيمة كانت افضل.

في هذا التحليل قيمة الفا = 0.786 وهي جيدة.



ملحق1: جداول التوزيعات الاحتمالية.

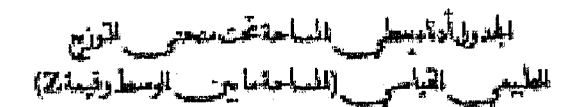
جدول التوزيع الطبيعي. Z

جدول تـوزيع ت. t

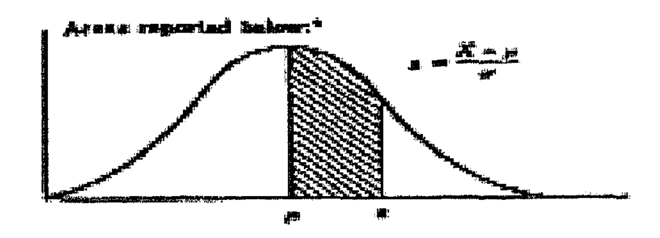
جدول تـوزيع ف. F

جدول تـوزيع كا^2. χ^2

جدول التوزيع الطبيعي Z.



الجدول أ: المساحة تحت منحنى التوزيع
الطبيعي القياسي (المساحة ما بين الوسط وقيمة Z)



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	2.612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4014
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4983	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987									
3.1	.4997									
3.2	.4999									

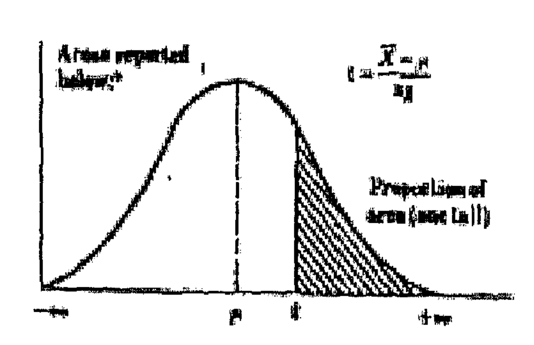
* Example: For z= 1.96, shaded area is 0.4750 out of the total area of 1.0000.

جدول توزيع T

الجدول أدناه يعطي قيمة t ≥

المقابلة للمساحة المظللة على يمينها α

Proportions of Area for the t Distributions



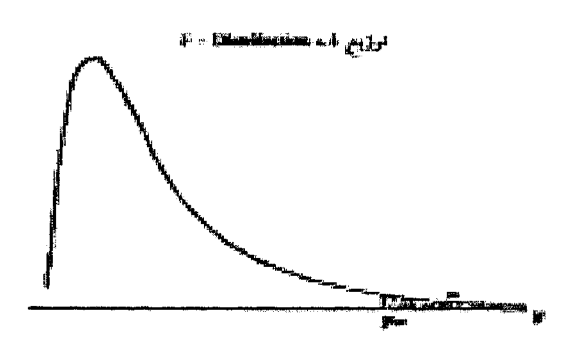
df	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898

df	0.10	0.05	0.025	0.01	0.005
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

* Example : For the shaded area to represent 0.05 of the area of 1.0,value t with 10 degrees of freedom is 1.812.

Source: From Table III of Fisher and Yates, Statistical Tables for Biological, Agricultural and Medical Research, 6th., 1974, published by Longman Group Ltd., London (previously published by Oliver & Boyd, Edinburgh), by permission of the authors and publishers.

جدول توزيع F

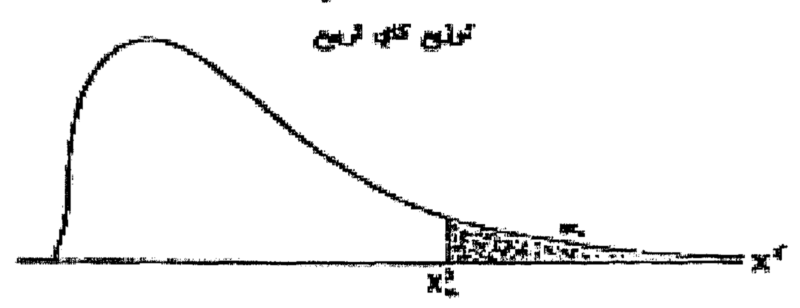


الصف الأول α = .05 والصف الثاني α = .01

مثال $F_{5,7}(.05)=3.97$ و $F_{5,7}(.01)=7.46$

درجات حرية المقام	درجات حرية البسط 1	2	3	4	5	6	7	8	9	10	11	12
1	161 4.052	200 4.999	216 5.403	225 5.625	230 5.764	234 5.859	237 5.928	239 5.981	241 6.022	242 6.056	243 6.082	244 6.106
2	18.51 98.49	19.00 99.00	19.16 99.17	19.25 99.25	19.30 99.30	19.33 99.33	19.36 99.36	19.37 99.37	19.38 99.39	19.39 99.40	19.40 99.41	19.41 99.42
3	10.13 34.12	9.55 30.82	9.28 29.46	9.12 28.71	9.01 28.24	8.94 27.91	8.88 27.67	8.84 27.49	8.81 27.34	8.78 27.23	8.76 27.13	8.74 27.05
4	7.71 21.20	6.94 18.00	6.59 16.69	6.39 15.98	6.26 15.52	6.16 15.21	6.09 14.98	6.04 14.80	6.00 14.66	5.96 14.54	5.93 14.45	5.91 14.37
5	6.61 16.26	5.79 13.27	5.41 12.06	5.19 11.39	5.05 10.97	4.95 10.67	4.88 10.45	4.82 10.29	4.78 10.15	4.74 10.05	4.70 9.96	4.68 9.89
6	5.99 13.74	5.14 10.92	4.76 9.78	4.53 9.15	4.39 8.75	4.28 8.47	4.21 8.26	4.15 8.10	4.10 7.98	4.06 7.87	4.03 7.79	4.00 7.72
7	5.59 12.25	4.74 9.55	4.35 8.45	4.12 7.85	3.97 7.46	3.87 7.19	3.79 7.00	3.73 6.84	3.68 6.71	3.63 6.62	3.60 6.54	3.57 6.47
8	5.32 11.26	4.46 8.65	4.07 7.59	3.84 7.01	3.69 6.63	3.58 6.37	3.50 6.19	3.44 6.03	3.39 5.91	3.34 5.82	3.31 5.74	3.28 5.67
9	5.12 10.56	4.26 8.02	3.86 6.99	3.63 6.42	3.48 6.06	3.37 5.80	3.29 5.62	3.23 5.47	3.18 5.35	3.13 5.26	3.10 5.18	3.07 5.11
10	4.96 10.04	4.10 7.56	3.71 6.55	3.48 5.99	3.33 5.64	3.22 5.39	3.14 5.21	3.07 5.06	3.02 4.95	2.97 4.85	2.94 4.78	2.91 4.71
11	4.84 9.65	3.98 7.20	3.59 6.22	3.36 5.67	3.20 5.32	3.09 5.07	3.01 4.88	2.95 4.74	2.90 4.63	2.86 4.54	2.82 4.46	2.79 4.40
12	4.75 9.33	3.88 6.93	3.49 5.95	3.26 5.41	3.11 5.06	3.00 4.82	2.92 4.65	2.85 4.50	2.80 4.39	2.76 4.30	2.72 4.22	2.69 4.16
13	4.67 9.07	3.80 6.70	3.41 5.74	3.18 5.20	3.02 4.86	2.92 4.62	2.84 4.44	2.77 4.30	2.72 4.19	2.67 4.10	2.63 4.02	2.60 3.96
14	4.60 8.86	3.74 6.51	3.34 5.56	3.11 5.03	2.96 4.60	2.85 4.46	2.77 4.28	2.70 4.14	2.65 4.03	2.60 3.94	2.56 3.86	2.53 3.80
15	4.54 8.68	3.68 6.36	3.29 5.42	3.06 4.89	2.90 4.56	2.79 4.32	2.70 4.14	2.64 4.00	2.59 3.89	2.55 3.80	2.51 3.73	2.48 3.67
16	4.49 8.53	3.63 6.23	3.24 5.29	3.01 4.77	2.85 4.44	2.74 4.20	2.66 4.03	2.59 3.89	2.54 3.78	2.49 3.69	2.45 3.61	2.42 3.55
17	4.45 8.40	3.59 6.11	3.20 5.18	2.96 4.67	2.81 4.34	2.70 4.10	2.62 3.93	2.55 3.79	2.50 3.68	2.45 3.59	2.41 3.52	2.38 3.45
18	4.41 8.28	3.55 6.01	3.16 5.09	2.93 4.28	2.77 4.25	2.66 4.01	2.58 3.85	2.51 3.71	2.46 3.60	2.41 3.51	2.37 3.44	2.34 2.37

جدول توزيع X^2



الجدول أدناه يعطي قيمة x^2_α المقابلة للمساحة المظللة وقيمتها α

درجات الحرية	المساحة المظللة						
	.99	.98	.95	.90	.80	.70	.50
1	.03157	.03628	.00393	0.0158	.0642	.148	.455
2	.0201	.0404	.103	.211	.446	.713	1.386
3	.115	.185	.352	.584	1.005	1.424	2.366
4	.297	.429	.711	1.064	1.649	2.195	3.357
5	.554	.752	1.145	1.610	2.343	3.000	4.351
6	.872	1.134	1.635	2.204	3.070	3.828	5.348
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344
9	2.088	2.532	3.325	4.168	5.390	6.393	8.343
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336

المصادر
Bibliography

المصـــادر العربيـــة Arabic Biblography

1. باهي، مصطفى حسين، عبد الفتاح احمد (2006). الاحصاء التطبيقي باستخدام الحزم الجاهزة SPSS, SAT . مصر، القاهرة: مكتبة الانجلو المصرية.
2. البلداوي، عبد الحميد (1997). *الإحصاء للعلوم الإدارية والتطبيقية*. الأردن، عمان: دار الشروق للنشر والتوزيع.
3. البـــياتي، محمـــود (2005). *تحليل البيانات الإحصائية باستخدام البرنامج الإحصائي SPSS*. الأردن، عمان: دار الحامد للنشر والتوزيع.
4. ابـــو حطـــب، فؤاد (1962). **القدرات العقلية** (ط5). مصر، القاهرة: مكتبة الانجلو المصرية.
5. الزعبي، بلال محمد والطلافحة، عباس (2000). *النظام الإحصائي SPSS فهم وتحليل البيانات الإحصائية*. الأردن، عمان: دار وائل للطباعة والنشر.
6. الـــسيد، فؤاد البهي (1987). **علم النفس الاحصائي وقياس العقل البشري**. مصر، القاهرة: دار الفكر العربي.
7. عـــبد الجـــبار، توفـــيق (1983). *التحليل الإحصائي في البحوث التربوية والنفسية والاجتماعـــية-الطـــرق اللامعلمية*. الكويت: مؤسسة الكويت للتقدم العلمي، إدارة التأليف والترجمة.
8. عـــبد الخالق، احمد محمد (1987). **الابعاد الاساسية للشخصية**. مصر، الاسكندرية: دار المعرفة الجامعية.
9. عـــلام، صـــلاح الـــدين (1993). *الأســـاليب الإحـــصائية الاستدلالية البارامترية واللابارامتـــرية في تحليل بيانات البحوث النفسية والتربوية*. مصر، القاهرة: دار الفكر العربي.

10. عــودة، احمد سليمان والخليلي، خليل (1988). ***الإحصاء للباحث في التربية والعلوم الإنسانية***. الأردن، عمان: دار الفرقان للنشر والتوزيع.

11. فتح الله، سعيد حسين (1988). ***مبادىء علم الإحصاء والطرق الإحصائية***. الأردن، المفرق: الأكاديمية.

12. فــراج، محمــد انور (2002). **المكونات العاملية للتفكير الناقد لدى طلاب كليات التــربية في ضــوء بعض المتغيرات**، رسالة دكتوراه غير منشورة، كلية التربية، جامعة الاسكندرية، الاسكندرية، مصر.

13. القاســم، محمــد علي (1987). ***أساليب الإحصاء التطبيقي***. الكويت: المعهد العربي للتخطيط بالكويت، دار الشباب للنشر والترجمة.

14. الكــيلاني، عــبد الله وعدس، عبد الرحمن (1986). ***القياس والتقويم في علم النفس والتــربية***. ترجمة: روبرت ثورندايك واليزابيث هيجن، الأردن، عمان: مركز الكتب الأردني.

15. منصور عوض ، عزام صبري ، علي قوقزة (1999). ***علم الإحصاء الوصفي المبرمج***. الأردن، عمان: دار صفاء للنشر.

16. موراي، ر.شبيرجل (1977). ***سلسلة ملخصات سشوم***، دار ماجدوهيل للنشر.

17. المنيــزل، عــبد الله فــلاح (2006). ***الإحــصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية SPSS***، الأردن، عمان: دار وائل للطباعة والنشر والتوزيع.

18. المنيــزل، عــبد الله والغرايبة، عايش (1995). ***الإحصاء التربوي تطبيقات باستخدام الرزم الإحصائية للعلوم الاجتماعية***.

19. الــنجار، نبيل جمعه (2004). ***مهارات الحاسوب***، الأردن، اربد: عالم الكتب الحديث للنشر والتوزيع.

20. النجار، نبيل جمعه (2007). ***الاحصاء في التربية والعلوم الانسانية مع تطبيقات برمجية SPSS*** ، الأردن، عمان: دار الحامد للنشر والتوزيع.

21. الــنجار، نبيل جمعه (2010). ***القياس والتقويم مع تطبيقات برمجية SPSS*** ، الأردن، عمان: دار الحامد للنشر والتوزيع.

22. النجار، فايزجمعه، النجار، نبيل جمعه والزعبي، ماجد راضي (2013). ***اساليب البحث العلمي:منظور تطبيقي***، الأردن، عمان: دار الحامد للنشر والتوزيع.

23. الهانـــسي، مختار محمود (1991). ***مقدمة في طرق التحليل الإحصائي***. قسم الإحصاء والرياضة والتأمين - كلية التجارة - جامعة الإسكندرية.

24. الهانـــسي، مختار محمود (1984). ***مقدمة في الإحصاء التحليلي***. بيروت: دار النهضة العربية للطباعة والنشر.

25. هكـــسي، تشارلز (1984). ***المفاهيم الأساسية في تصميم التجارب***. تعريب: خماس، قيس سبع، العراق، بغداد: الجامعة المستنصرية.

English Biblography المصادر الأجنبية

1. Albert K. Kurtz, Samuel T. Mayo (1979). ***Statistical Methods in Education and Psychology***. New York: Springer-Verlag, New York Inc.
2. Andy, Field (2005). ***Discovering Statistics Using SPSS***. (2nd ed.). London: Sage Publications Ltd, ECIY 1SP.
3. Berenson, L. Mark & David M. Levine. (1992). ***Basic Business Statistics Concepts and Applications***, (5th ed.).
4. Bobko, Philip (2001), ***Correlation and regression***, (2nd ed.). Thousand Oaks, CA: Sage Oublications. Introductory text which Includes coverage of range restriction, trivariate correlation.
5. Chen, P. Y. and P. M. Popovich (2002). ***Correlation: Parametric and nnparametric measures***. Thousand Oaks, CA: Sage Publications.
6. Cochran, W. G. (1997). ***Sampling Techniques*** (3rd ed.). New York.
7. Cohen, Jacob (1988). ***Statistical Power Analysis for the Behavioral Sciences***. Hillsdale, NJ: Lawrence Erlbaum Assoc; ISBN: 0805802835.
8. Crawley, M. J. (2005). ***Statistics An Introduction Using R***. England: John Wiley & Sons Ltd, West Sussex.
9. Crocker, Lind, and Algina, Janes .(1986). ***Introduction to Classical and Modern Test Theory***. N. Y: Holt, Rineder and Winson
10. David, R. Anderson, Dennis J. Sweeney , Thomas A. Williams (2002). ***Essentials of Modern Business Statistics with Microsoft Excel*** (2nd ed.). South-Western Educational Publishing.

11. David, R. Anderson, Dennis, J. Sweeney , Thomas A. Williams (2006). ***Essentials of Modern Business Statistics*** (3rd ed.). South-Western College Publishing.
12. David, R. Anderson, Dennis, J. Sweeney , Thomas, A. Williams (2005). ***Statistics for Business and Economics*** (9th ed.). South-Western College Publishing.
13. David Anderson, Dennis J. Sweeney & Thomas Williams, (1981). ***Introduction to Statistics***, West Publishing Co.
14. Daniel, Wayne W. (1990). ***Applied Nonparametric Statistics,*** (2nd ed.). PWS-Kent Publishing Company-Boston.
15. Daniel, Wayne W. (1995). ***Biostatistics A Foundation for Analysis in the Health Sciences*** (6th ed.). London: John Wiley and Sons Inc.
16. Deming, W. (1982). ***Applied Regression Analysis***. London: John Wiley and Sons Inc.
17. Dominick Salvator (1982). ***Theory and Problems of Statistics and Econometrics***, McGraw-Hill Book Co.
18. Draper N. & Smith H. (1990). ***Applied Regression Analysis***. London: John Wiley and Sons Inc.
19. Guilford, J.P (1961). **Factorial analysis to Psychology.**
20. Glass, G. & Hopkins, K. D.(1984). ***Statistical Methods in Educational and psychology(2nd ed.).*** Englewood Cliffs, New Jersey: Prentice-Hall.
21. Gary W. Heiman. (1992). ***Basic Statistics for the Behavioral Sciences***. Houghton Mifflin Company, Boston, U.S.A
22. Graber, S. B., Kristin E. Voelki, T. W. & Others .(1997). ***SPSS Guide to the New Statistical Analysis of Data***, New York: Springer.
23. Grant, E. & Leavenworth R., (1980). ***Statistical Quality Control,*** (5th ed.). New York: McGraw-Hill.
24. Gravetter, F. O.(1988). ***Statistics for the behavioral sciences.*** New York: West Publishing company.
25. Green, Samuel B. & Niel J. Salkind (1997). ***Using SPSS for Windows: Analyzing and Understanding Data,*** Upper Saddle River. NJ: Prentice Hall.
26. Gregory, B., (1991). ***Introduction to Quality Management Assurance and Control,*** (Macmillan International ed). New York: Maxwell.
27. Hays, W. L. (1980). ***Statistics for Social Sciences*** (3rd ed.). New York: Holt Rinehart and Winston.

28. Hays, W. L. (1989). ***Statistics*** (4th ed.). New York: Holt Rinehart and Winston.

29. Hogg, N. R. & Carag, T. A. (1995). ***Introduction to Mathematical Statistics*** (5th ed.).Prentice-Hall Inc.

30. Howitt, Dennis & Duncan, Cramer (1996). ***A Guide to Computing Statistical with SPSS for Windows,*** New York: Prentice-Hall/ Harvester Wheatsheaf.

31. Jaynes, E. T. & G. Larry Bretthorst (2003). ***Probability Theory: The Logic of Science***. Cambridge University.

32. Johnson, Richard A. & Wichern, D. (1992). Applied Multivariate Statistical (3rd). New jersey: prencac-hall.

33. Juran, J. & Gryna, F. (1970). ***Quality Planning and Analysis***. New Yourk: McGraw-Hill.

34. Kendall, Maurice and Jean Dickinson Gibbons (1990). ***Rank Correlation Methods***. (5th ed.). NY: Oxford Univ Press; ISBN: 0195208374.

35. Kenneth, D. Hopkins, Julian C. Stanley & B. R. Hopkins .(1995). ***Educational and Psychological Measurement and Evaluation.*** (7th ed.). Allyn & Bacon, Needham Hights, USA.

36. Larson, Harold. J. (1982). ***Introduction to Probability Theory and Statistical Inference*** (3rd ed.). John Wiely & Sons Inc.

37. Larson, Harold. J. (1995). ***Introduction to Probability,*** Addison Wesley Publishing Company Inc.

38. Larry, J. Stephens (2005). ***Schaum's Outline of Beginning Statistics*** (2nd ed.). McGraw-Hill.

39. Mark, S. Aldenderfer & Roger, K. Blashfield (1991). ***Cluster Analysis*** (8th ed.). sage Publications - The International Professional Publishers, Eighth Printing.

40. (Manual) (1997). ***SPSS Base 7.5 Application Guide***, SPSS Inc.

41. (Manual) (1997). ***SPSS Base 7.5 for Windows User Guide***, SPSS Inc.

42. (Manual) (1994). ***SPSS Advanced Statistics 6.1***, SPSS Inc.

43. Marija J. Norusis (1993). ***SPSS for Windows Base System User's Guide Release 6.0 (Manual)***, SPSS Inc.

44. McClave, J. T. & Benson, G. P. (1991). ***Statistics For Business and Economics*** (5th ed.). San Fransisco Dellan.

45. Mendenhall W. & Sincich (1991). ***Statistics For Engineering and the Sciences*** (3rd ed.). New York: MaCmillan Publishing Co.

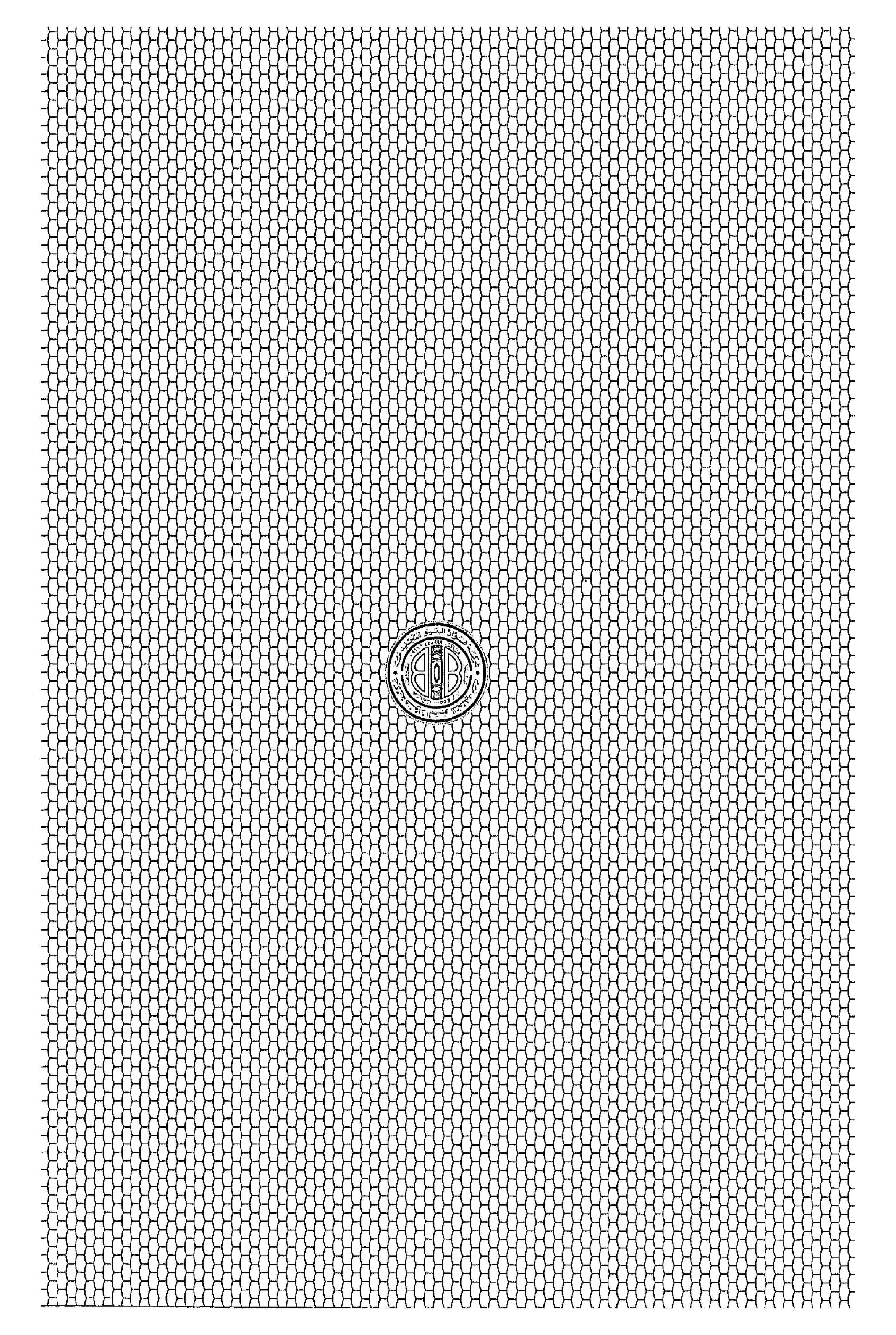
46. Minium, E. W.(1978). ***Statistical Reasoning in psychology and Education*** (2nd ed.). New York: John Wiley and Sons.

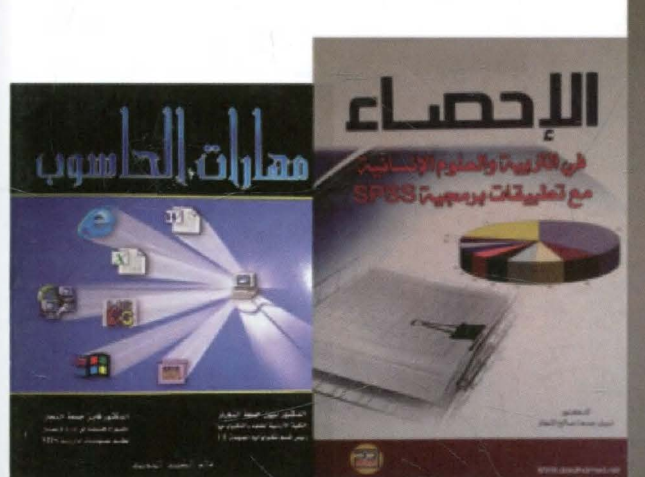
47. Morris, H. (1983). ***Statistical Analysis for Decision Making.*** (3rd ed.). USA: Harcourt Brace Jovanovish Inc.

48. Nelile, A. & Kennedy, J. (1974). ***Basic Statistical Methods for Engineers & Scientists,*** London: Inter Text Books.

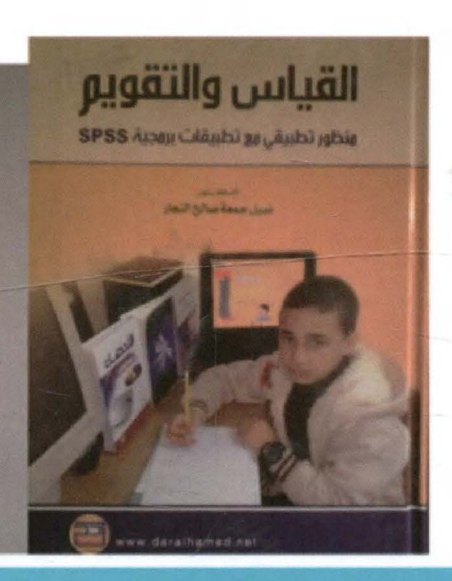
49. Parzen, E. (1960). ***Modern Probability theory and its application,*** New York: John Wiley and Sons.

50. Siegel, S. (1956). ***Nonparametric statistics for the behavioral sciences.*** NY: McGraw-Hill.









الإحصاء التحليلي

مع تطبيقات برمجية SPSS



دار الحامد للنشر والتوزيع

الأردن-عمان

هاتف: 5231081 فاكس: +96265235594

ص.ب:366 عمان 11941 الأردن

E-mail:dar_alhamed@hotmail.com

E-mail:Daralhamed@yahoo.com

